Durham E-Theses

Ocular higher-order aberrations and visual performance

YOUNG, LAURA,KATE

How to cite:

YOUNG, LAURA,KATE (2011) *Ocular higher-order aberrations and visual performance*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/3282/

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, Durham University, University Office, Old Elvet, Durham DH1 3HP
e-mail: e-theses.admin@dur.ac.uk Tel: +44 0191 334 6107
http://etheses.dur.ac.uk

Ocular higher-order aberrations and visual performance

Laura K. Young

A Thesis presented for the degree of
Doctor of Philosophy

Centre for Advanced Instrumentation

Department of Physics

Durham University

England

October 2011

Abstract

Since adaptive optics was first used to correct the monochromatic aberrations of the eye over a decade ago there has been considerable interest in correcting the ocular aberrations beyond defocus and astigmatism. In order to understand the prospective benefits of correcting these higher-order aberrations it is important to study their effect on visual performance. From a clinical perspective it is important to know how different types of aberration can affect visual performance so that wavefront measurements can be better understood. Visual performance is determined by a combination of optical and neural factors. It is important to consider how degradations in the optical quality of the eye can impact the neural processes involved in visual tasks such as object recognition.

In this thesis we present a study of the effects of three types of aberration, defocus, coma and secondary astigmatism, on letter recognition and reading performance. In the course of this work we also characterise the repeatability of the Zywave aberrometer, which we used to measure our subjects' ocular wavefronts. We use stimuli that have these aberrations applied in their rendering to examine the differences between these aberrations and how they differ with respect to the visual task. We find that secondary astigmatism causes the largest impairment to both letter recognition and reading performance, followed by defocus. Coma causes comparatively smaller degradations to performance but its effect is different depending on the visual task. We can predict the reduction in performance based on a simple cross-correlation model of letter confusability. The relationship between these predictions and the experimental results are the same for all three aberrations, in the case of single letter recognition. In reading however, the relationship is different for coma. We suggest that coma causes lateral masking effects and may additionally disrupt the planning of eye movements. Coma slows reading, but does not specifically impair word identification whereas defocus and secondary astigmatism do. We attribute disruptions in word identification to the dramatic effects defocus and secondary astigmatism have on the form of a letter.

Declaration

The work in this thesis is based on research carried out in the Centre for Advanced Instrumentation, Department of Physics and in the Department of Psychology at the University of Durham, England. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is the sole work of the author unless referenced to the contrary in the text. My supervisors, Dr. Gordon Love, Dr. Richard Myers and Dr. Hannah Smithson, and I received specialist guidance on interpretation of the results of the reading experiment in Chapter 7 from our collaborator, Prof. Simon Liversedge, in the School of Psychology at the University of Southampton. Work that involved human subjects was conducted in accordance with the declaration of Helsinki and the experimental protocol was reviewed and approved by the local ethics committee in the Department of Psychology.

Publications

- Young, L. K., Liversedge, S. P., Myers, R. M., Love, G. D., & Smithson, H. E. (2011). Not all aberrations are equal: Reading impairment depends on aberration type and magnitude. *Journal of Vision*, 11(13):20, 1-19.

- Guzmán, D., Javier de Cos Juez, F., Lasheras, F. S., Myers, R. and Young, L. (2010). Deformable mirror model for open-loop adaptive optics using multivariate adaptive regression splines, *Optics Express*, 18(7):64926505.

- Guzmán, D., Javier De Cos Juez, F., Myers, R., Lasheras, F. S., Young, L. K. and Guesalaga, A. (2010). Deformable mirror models for open-loop adaptive optics using non-parametric estimation techniques, *Proc. SPIE*, 7736:77361C.

- Gendron, E., Morris, T., Hubert, Z., Myers, R., Longmore, A., Rousset, G., Talbot, G., Vidal, F., Dipper, N., Gratadour, D., Looker, N., Brangier, M., Younger, E., Sevin, A., Basden, A., Perret, D., Young, L., Atkinson, D., Chemla, F., Henry, D., Butterley,

T., Laporte, P., Guzman, D., Marteaud, M., Geng, D., Vedrenne, N., Harrison, M., Fusco, T., Guesalaga, A., Dunlop, C., Todd, S., Dee, K., Dickson, C., Greenaway, A., Stobie, B., Dalgarno, H. and Skvarc, L. (2010). Status update of the Canary on-sky MOAO demonstrator, *Proc. SPIE*, 7736:77360P.

- Myers, R. M., Hubert, Z., Morris, T. J., Gendron, E., Dipper, N. A., Kellerer, A., Goodsell, S. J., Rousset, G., Younger, E., Marteaud, M., Basden, A. G., Chemla, F., Guzman, C. D., Fusco, T., Geng, D., Le Roux, B., Harrison, M. A., Longmore, A. J., Young, L. K., Vidal, F. and Greenaway, A. H., (2008). CANARY: The on-sky NGS/LGS MOAO demonstrator for EAGLE, *Proc. SPIE*, 7015:70150E.

- Guzman, D., Guesalaga, A., Myers, R., Sharples, R., Morris, T., Basden, A., Saunter, C., Dipper, N., Young, L., Rodriguez, L., Reyes, M. and Martin, Y. (2008) Deformable mirror controller for open-loop adaptive optics, *Proc. SPIE*, 7015:70153X.

Copyright ©2011 Laura Young. "The copyright of this thesis rests with the author. No quotation from it should be published without the prior written consent and information derived from it should be acknowledged".

Acknowledgements

Firstly, I would like to thank my supervisors Gordon Love, Hannah Smithson and Richard Myers for the opportunity to work on such an interesting project and kindling my passion for studying vision. Their support and guidance has been invaluable over the years. I would also like to thank Simon Liversedge at the University of Southampton for his advice on the reading experiment. To all four I extend my gratitude for their encouragement when submission of the notorious JoV paper seemed a never-ending task.

I would like to acknowledge funding from a Science and Technologies Facilities Council postgraduate studentship.

To the other members of the Centre for Advanced Instrumentation, and couple of people from up the hill: Thank you for both the useful discussions and for the non-useful, but none-the-less entertaining conversations. I'd also like to say thank you for your participation in my experiments, without which I would have no results! Particular thanks go to Chris, for having a glass of water on hand (even when I didn't want it), to Andy for forcing me up the climbing wall (even when I didn't feel like it) and to Mike, Graham, James and all of the others who have had a beer on hand (even when I probably didn't need it). To Tim: Thank you for being there, you've kept me sane, which is a formidable feat. Finally I'd like to thank my family for their unending support and encouragement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Vision is our dominant sense with around 27% of the brain being devoted to visual function (Van Essen, 2004). Vision is more than just an imaging task for we must not only be able to extract behaviourally relevant information from an image but also form a representation of the object in view. Recognition of an object is a difficult problem since the image on the retina can vary dramatically due factors such as viewing angle, distance, illumination, motion, occlusion or even superficial changes in the shape of the object. We are able to hold a representation of an object in such a way that it can be recognised in spite of these differences, which are not intrinsic to the object. As an example, we are capable of recognising letters and words despite differences in font and even differences in handwriting (see Figure 1.1 for example). In fact the interpretation of identical ambiguous letters is possible with the aid of context, as shown in Figure 1.2. Subjectively, letter identification appears to be achieved nearly instantaneously with little effort. The first stage of this process occurs in the retina, which is an extension of the brain connected by the optic nerve. In peripheral retina the spatial resolution is limited by the sampling rate of the photoreceptors and subsequently by the neural sampling rate imposed by successive layers of processing cells. In the fovea spatial resolution is thought to be limited by the retinal image quality, which is determined by the optical properties of the eye.

(a) (b)

Figure. 1.1: Example of skewing effect on the letter 'A' (shown in black) showing a) the original letter 'A' overlaid on a letter 'R' (shown in white) and b) a skewed letter 'A' over laid on a letter 'R'. Despite the differences between the appearance of the letter 'A' in panels a) and b) and their overlap with the letter 'R' we still perceive both black letters as an 'A'.

Figure. 1.2: Example of how context can influence letter and word identification. One finds it difficult to read anything other than "the cat" despite the ambiguous character in the centre of each word.

The optical components of the eye are imperfect and cause distortions in the retinal image. Johannes Kepler (1604) was first to describe the possibility of correcting myopia or hyperopia using concave or convex lenses respectively. Almost 200 years later Thomas Young (1801) recorded the existence of astigmatism in his own eye. Herman von Helmholtz (1867) was the first to suggest that the human eye suffered from monochromatic aberrations beyond defocus and astigmatism and that these higher-order aberrations could not be corrected with a simple lens. Measurement of these aberrations was first performed by Smirnov (1961) and later Liang *et al.* (1994) measured them with a Shack-Hartmann wavefront sensor. For the first time correction of the higher-order aberrations in the eye was achieved using an adaptive optics (AO) system (Liang & Williams, 1997; Liang, Williams, & Miller, 1997). Although the motivation for this system was initially to record high resolution retinal images, an improvement in contrast sensitivity was observed (Liang *et al.*, 1997). Since then development of AO systems for studying the eye and visual function has

progressed, incorporating real-time correction (Hofer, Artal, Singer, Aragón, & Williams, 2001; Fernández, Iglesias, & Artal, 2001; Diaz-Santana & Dainty, 2003), binocular correction (Fernández, Prieto, & Artal, 2009) and AO corrected, targeted and stabilised stimulus presentation (Arathorn *et al.*, 2007). The use of AO for vision science falls into two categories; those used for high resolution retinal imaging and those used for studying visual function. In this thesis we are concerned with the role of ocular aberrations in visual performance so we will focus the latter of these two categories.

Early work testing visual function has demonstrated the benefits of using an AO system to correct for the optical defects of the eye (see Chapter 2 for a summary). The prospect of correcting down to the diffraction limit is an exciting one, having both clinical and research applications. In particular, improving visual performance by correcting the higher-order aberrations of the eye with a portable device such as a phase plate (Yoon, Jeong, Cox, & Williams, 2004) or contact lens (López-Gil, Chateau, Castejón-Monchón, Artal, & Benito, 2003) has been studied. The feasibility of correcting higher-order aberrations with wavefront-guided laser refractive surgery is still unclear (Netto, Dupps, & Wilson, 2006). This procedure often introduces higher-order aberrations and this is difficult to prevent in living tissue due to unpredictable variations in wound healing. In short, while it is possible provide a static, open-loop correction of the higher-order aberrations of the eye this is difficult and so it is important to understand the potential benefits. In particular we need to know how well we can measure the eye's aberrations and how significant the errors on that measurement are in terms of visual performance. In the case of refractive surgery, we also wish to know the impact of inducing aberrations in the ablation process. In this thesis we study the effects of different types of aberration on visual performance to aid the interpretation of wavefront measurements in terms of real life tasks. Applegate *et al.* (2002) have previously used a letter chart to study the effect on visual acuity of inducing aberrations. We do not often need to recognise single letters on their own but we do use letter recognition in reading, which is a common visual task. Reading is a time-critical task

that, in the presence of an aberration, subjects may be capable of completing but with a significant increase in difficulty. Additionally, by studying reading we can assess the impact of aberrations on a higher-level process involving language. It is one thing to recognise a single letter, or even a word, but it is a completely different matter to extract contextual information and meaning. Aberrations that interfere with letter recognition will undoubtedly have far reaching effects on language processing, which relies on efficient identification of letters and words. We will show that the effect of certain monochromatic aberrations on visual performance is task-dependent. Moreover, the differences between the process of letter identification and that of reading are significant and we cannot take a measure of letter acuity as a standard description of the effect of an aberration on vision.

1.1. Synopsis

This thesis is concerned with the impact of monochromatic aberrations on reading performance. We also examined their impact on letter identification, in terms of contrast threshold, as a comparison. We imposed these aberrations in the rendering of the stimuli, rather than optically. In order to understand the implications of this we studied the effect of subjects' own aberrations on the retinal image. This involved repeated measurements using a commercial Shack-Hartmann wavefront sensor, which we tested for repeatability. This thesis is presented in the following way:

- Chapter 2 summarises the aberration structure of the eye and discusses Zernike polynomials, which are the standard basis set for describing an ocular wavefront measurement. We introduce the concepts of the point spread function, the optical transfer function and the visual Strehl ratio, which are important concepts in studying image quality. We explain convolution and cross-correlation, which are used in this thesis for stimulus generation and for modelling letter recognition respectively.

- Chapter 3 describes the structure of the visual system from photoreceptors in the

retina to neurons in the visual cortex. We introduce the idea of a receptive field and explain how receptive fields can process spatial frequency and spatial phase. We outline the interactive activation model of word recognition, which is a popular model for describing the processes of letter and word recognition. We give a summary of the use of eye movement data to study reading performance and explain some important concepts that are referred to in Chapter 7.

- Chapter 4 concerns the simulation of ocular aberrations and explains the method of rendering stimuli with an aberration. We discuss the implications of inducing an aberration in this way rather than using an AO system and we form justifications for this methodology. We conclude that the effect of subjects' ocular aberrations on the retinal image was minimal in our experiments. We describe a prediction of the effect different types of aberration might have on performance. We analyse the effect of different types of aberration on the distinguishability of letters by comparing rendered letters via a cross-correlation. We also study the effect of these aberrations on relative changes in letter position, which is important for reading. Clinical considerations are also noted. It is not possible to test every type and combination of aberrations so we chose to study only three. Based on these analyses we chose to study defocus, coma and secondary astigmatism.

- Chapter 5 describes a repeatability study of the Zywave aberrometer. We explain how the Zywave works and summarise previous studies on its repeatability. We discuss the sources of error in the ocular aberration measurement and quantify these based on measurements taken with a model eye. Having characterised the system-induced aberrations we detail repeated measurements taken in human eyes over the course of a week. From these results we conclude that the Zywave measurements may not be repeatable for certain Zernike modes and that this is likely due to the small amplitudes of these modes in both the model and human eyes. The largest sources of error are attributed to changes in pupil size, changes in focal distance and alignment errors. In

human eyes we find that the aberration measurement does not change significantly from day to day within the limits of the instruments' repeatability. There were small, but significant, variations during the day in one subject suggesting that the stability of ocular aberrations may vary in the normal population. We conclude that repeated aberration measurements should be made prior to testing visual performance to get the best estimate of a subject's ocular wavefront and differences in pupil radius should be accounted for.

- Chapter 6 presents a study of the effects of defocus, coma and secondary astigmatism on the contrast threshold for letter recognition. We describe a maximum likelihood staircase algorithm, which is the process we use to measure contrast threshold. We evaluate the expected performance of this algorithm and, based on this, we set criteria for determining its successful convergence on the appropriate threshold value. The experimental results and subsequent analysis are discussed and we make comparisons to the predictions made by the cross-correlation analysis described in Chapter 4. We show that coma has a smaller effect the contrast threshold for letter recognition than defocus or secondary astigmatism and conclude that this is due to the feature-preserving nature of coma. We demonstrate the ability to predict increases in contrast threshold caused these aberrations with a single relationship to the confusability metric determined by our cross-correlation model.

- Chapter 7 presents a study of the effects of defocus, coma and secondary astigmatism on reading performance. We describe global measures of reading performance, such as the average fixation time on a word calculated over a whole sentence, which indicate a more general effect on reading. Local word-specific effects are also revealed by embedding a target word in the sentence. This target word has either a high or low lexical frequency, which is an indication of how common the word is in language and forms an index for the ease with which a word is identified. Less common words are identified less easily and so we expect to measure a lexical frequency effect with

subjects fixating longer on these words (Rayner, Liversedge, & White, 2005; Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Rayner, 1998). We find that defocus and secondary astigmatism cause an increase in the size of the lexical frequency effect with increasing amplitude of aberration for the total fixation time on the target word. This suggests that these aberrations degrade the retinal image quality in such a way that subsequent recognition processes are increasingly impaired as the amplitude increases. We propose that this provides strong evidence that these two types of aberration impair performance by degrading the form of letters. We compare these results to our prediction of performance based on the confusion metric and find that while defocus and secondary astigmatism can be predicted with the same relationship, coma cannot. We suggest that this indicates a different source of performance loss when reading with coma. We attribute this to lateral masking effects and disruptions to saccade planning, neither of which occur when letters are presented in isolation.

- Chapter 8 concludes with a summary of the findings of this thesis and gives suggestions for future work. These include improved modelling of the distinguishability of letters using techniques such as mutual information and further investigation using different types and combinations of aberrations. It would be of particular interest to study how the orientation of coma affects reading since the lateral masking effect may not present itself with vertical coma, where the letters are not 'smeared' into one another. In the case of horizontal coma we would also suggested testing the direction of this affect to determine whether 'smearing' from left to right is more or less detrimental than from right to left. Although lateral masking occurs in both directions, the perceptual span is not symmetric about fixation (e.g. McConkie & Rayner, 1976). Additionally a lack of masking of the first letter in a word may be beneficial as eye movements are biased towards the beginning of a word (McConkie, Kerr, Reddix, & Zola, 1988).

Chapter 2

Ocular aberrations

2.1. Introduction

The first limitation to visual acuity is that imposed by the optical properties of the eye. Any subsequent processing of the visual scene is limited by the retinal image quality. The purpose of this thesis is to examine how the optical characteristics of the eye can impact visual performance, specifically in relation to letter recognition and reading. The theory discussed in the next two chapters underpins the objectives of this work. This chapter focuses on the optical properties of the eye, specifically the nature of the aberrations of the eye and how they are quantified. As we will discuss, the wavefront measurement can be broken down into a series of orthogonal components. It is important to understand how these components differ in their affect on visual performance. This is necessary for the interpretation of ocular wavefront measurements and for understanding the potential benefits of vision correction. Chapter 3 will introduce the visual system and detail how it extracts visual information from the retinal image and supports object recognition from it. The perceptual processes described Chapter 3 fundamentally differentiate the implications of aberrations in the eye from those in other systems such as the atmosphere, for which adaptive optics was initially proposed (Babcock, 1953). It is crucial to consider how neural

encoding of degraded visual information may further affect perception.

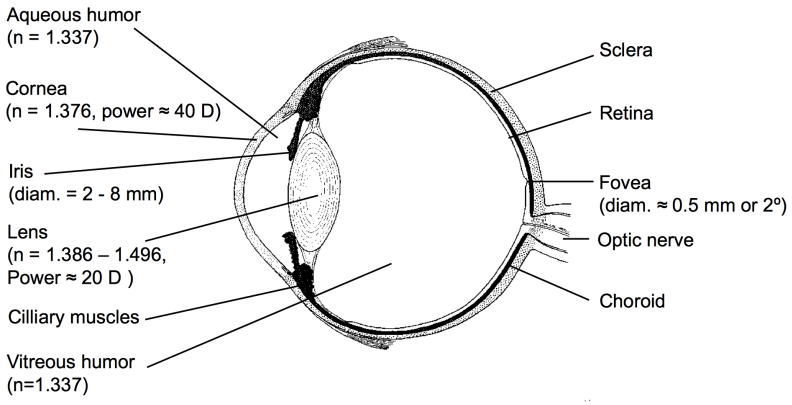

Figure. 2.1: Cross-section of the eye, which is approximately 24 cm long, showing its main components. There are five main refracting elements; the anterior and posterior surfaces of the cornea, the anterior and posterior surfaces of the lens and the refractive index gradient in the lens. The eye has a focal power of approximately 60 Diopters of which two thirds is contributed by the cornea and the remaining third by the lens. The average young human eye has an accommodation range of about 16 Diopters but this reduces with age.

2.2. Aberrations in the human eye

The structure of the human eye is given in Figure 2.1, showing its main components. There are four optical surfaces; the anterior and posterior surfaces of the cornea and the anterior and posterior surfaces of the lens. There is also a refractive index gradient across the lens that varies both axially and radially. Variations in the aqueous and vitreous fluids can also have an affect on the optical quality of the eye, as can changes in the tear film (Koh, Maeda, & Kuroda, 2002; Dubra, Paterson, & Dainty, 2004) and deformations in the structures of the eye, such as the retinal surface, due to blood flow (Hampson, Munro, Paterson, & Dainty, 2005). The iris acts as an aperture stop (a stop in the pupil plane) and so contributes to diffraction effects. The optical components of the eye are not perfect and the aberrations from each element contribute differently to the retinal image quality. In a small number of young subjects, Artal *et al.* (2001) found that the aberrations of the lens

partially compensated for those of the cornea. However, this complimentary distribution of the aberrations may not be optimal off-axis and so could contribute to the reduction in retinal image quality that is observed away from the fovea. It has been particularly noted that, like in other optical systems, coma increases significantly with retinal eccentricity (Navarro, Moreno, & Dorronsoro, 1998; Artal *et al.*, 2001; Atchison & Scott, 2002). In reality the optical quality of the eye is not likely to cause a significant limitation on peripheral visual acuity. The lower photoreceptor density and reduced neural sampling exhibited in peripheral retina are likely to be the most substantial constraint (Lundström, Manzanera, Prieto, & Ayala, 2007). However, the introduction of optical blur to peripheral retina has been shown to prevent aliasing between receptoral and post-receptoral spatial sampling by limiting the transmission of high spatial frequencies (Williams, Artal, Navarro, McMahon, & Brainard, 1996). Off-axis aberrations are particularly important in the context of this thesis since reading requires the extraction of visual information from parafoveal and peripheral retina, as we will discuss in the next chapter. Ocular aberrations change over time with a temporal frequency of 5-6 Hz due to processes such as accommodation (which particularly influences spherical aberration), changes in the tear film and ocular fluids and eye movements (Hofer *et al.*, 2001). In the longer term, aberrations are known to increase with age, particularly third order aberrations (McLellan, Marcos, & Burns, 2001; Artal, Berrio, Guirao, & Piers, 2002). The eye is additionally affected by chromatic aberrations, a comment on which will be made in Chapter 4.

2.2.1. Zernike polynomials

The normal human eye suffers from monochromatic aberrations that degrade the retinal image. These aberrations are described in terms of wavefront error, which is the deviation from a perfect plane wave. The wavefront can be broken down in to different types of basis set, such as Seidel aberrations for example, which describe aberrations in an optical system that is symmetric about a central optical axis. The standard method for describing

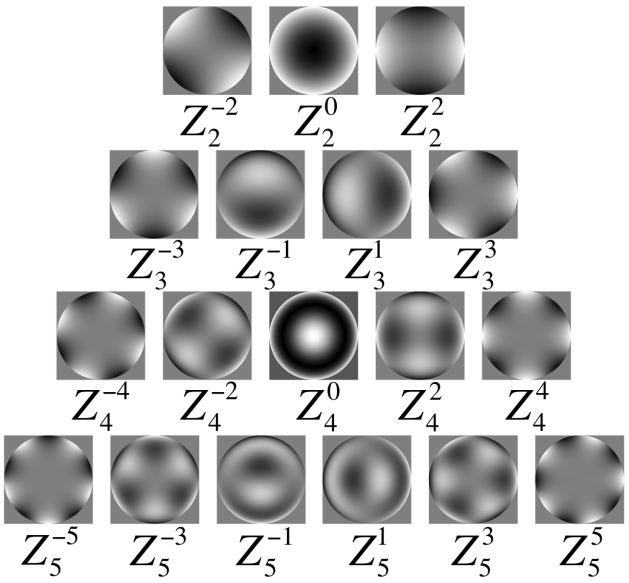

Figure. 2.2: The Zernike pyramid showing wavefront modes arranged in terms of their radial (n) and angular (m) orders labelled as Z_n^m. Higher-order aberrations are those with $n > 2$. The modes investigated in thesis are defocus (Z_2^0), coma (Z_3^1) and secondary astigmatism (Z_4^2).

these aberrations is to use Zernike coefficients (Thibos, Applegate, Schwiegerling, & Webb, 2000) that represent the contribution of their corresponding Zernike modes to the total wavefront error. Zernike polynomials (Zernike, 1934) are a basis set that are orthogonal over a unit circle so the contributions of different modes can be independently measured. The value of each coefficient is independent of the number of modes into which the wavefront is decomposed. Zernike polynomials are constructed in such a way that higher-order modes can be balanced by lower-order modes. Using a double indexing scheme they are grouped in terms of their radial frequency (the highest power of the radial component),n, and angular frequency (the angular frequency of the sinusoidal component), m. This grouping leads to the conventional Zernike pyramid shown in Figure 2.2. The value of n is a positive integer or zero and the value of m is constrained such that $-n \leq m \leq n$ in integer steps of 2. The

wave aberration is represented with a Zernike polynomial expansion:

$$W(r,\,\theta) = \sum_{n=0}^{n_{\max}} \sum_{m=-n}^{n} c_n^m Z_n^m(r,\,\theta), \tag{2.1}$$

where r is the radial coordinate, θ is the angular coordinate, c_n^m is the Zernike coefficient corresponding to the Zernike mode $Z_n^m(r,\,\theta)$ and $n_{\max}$ is the highest radial order to be considered. The polynomials describing these Zernike modes are defined by

$$Z_n^m(r,\,\theta) = \left\{ \begin{array}{ll} N_n^m R_n^{|m|}(r) \cos m\theta & for\ m \geq 0 \\ N_n^m R_n^{|m|}(r) \sin m\theta & for\ m < 0 \end{array} \right\}, \tag{2.2}$$

where $R_n^{|m|}(r)$ is the radial component given by

$$R_n^{|m|}(r) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)!}{s![0.5(n+|m|-s)]![0.5(n-|m|-s]!} r^{n-2s}. \tag{2.3}$$

The normalisation constant is given by

$$N_n^m = \sqrt{\frac{2(n+1)}{1+\delta_{m0}}}, \tag{2.4}$$

where δ_{m0} is the Kronecker delta function ($\delta_{m0} = 1$ for $m = 0$ and $\delta_{m0} = 0$ for $m \neq 0$). Although Zernike coefficients are the standard quantity for measuring aberrations in the eye, they are not necessarily suitable for describing visual performance. The same root mean square (rms) amplitude of two different Zernike modes can give significantly different retinal image qualities and more importantly can affect visual acuity by different amounts (see Applegate *et al.*, 2002, for example). In order to fully understand an ocular wavefront measurement it is necessary to characterise the relationship between the rms amplitude of a Zernike mode and the impact on visual performance, quantified by a metric specific to the visual task under consideration. The value of a Zernike coefficient is dependent on the radius of the pupil. Equivalent defocus, M_e, is an aberration metric that is independent

of pupil size and is defined as the amount of defocus in diopters that produces the same wavefront variance as a given aberration. It is determined using

$$M_e = \frac{4\pi\sqrt{3}\ RMS}{A} = \frac{4\sqrt{3}\ RMS}{r^2}, \tag{2.5}$$

where RMS is the wavefront variance, A is the area of the pupil and r is the pupil radius (Thibos, Hong, Bradley, & Cheng, 2002). The amplitudes of aberration used in this thesis are described in the relevant chapters and equivalent defocus values are also quoted. The size of subjects' pupils (or effective pupils) were not the same in the letter and reading experiments, as will be explained in Chapters 6 and 7. It is possible, using Equation 2.5 to approximately equate the amplitudes of aberration used for the different pupil sizes.

In this thesis we investigate three different Zernike modes in isolation. The justification for our choice of Zernike modes is given in Chapter 4. We have chosen to investigate defocus, coma and secondary astigmatism. Figure 2.3 shows ray diagrams for these aberrations and demonstrates the effects of each of these aberrations on the image. In simple terms we can think of defocus as causing a symmetric blur of the image, coma causing the image to smear in one direction and secondary astigmatism causing a shadowing effect akin to double vision. In subsequent chapters we will show the effects of these aberrations on letters and words.

2.2.2. Previous studies of correcting or inducing ocular aberrations

Smirnov (1961) was the first to measure the aberrations of the eye as a whole, commenting that it would be possible to correct them and recover the diffraction limited image quality. This was not realised until the first adaptive optics system for the eye was built by Liang and Williams (1997) (see also Liang *et al.*, 1997) based on the earlier work on a Shack-Hartmann wavefront sensor for the eye (Liang *et al.*, 1994). The basic architecture of an adaptive optics system comprises of a wavefront sensor (for example a Shack-Hartmann sensor such as the one described in Chapter 5), a wavefront corrector and a control algorithm.

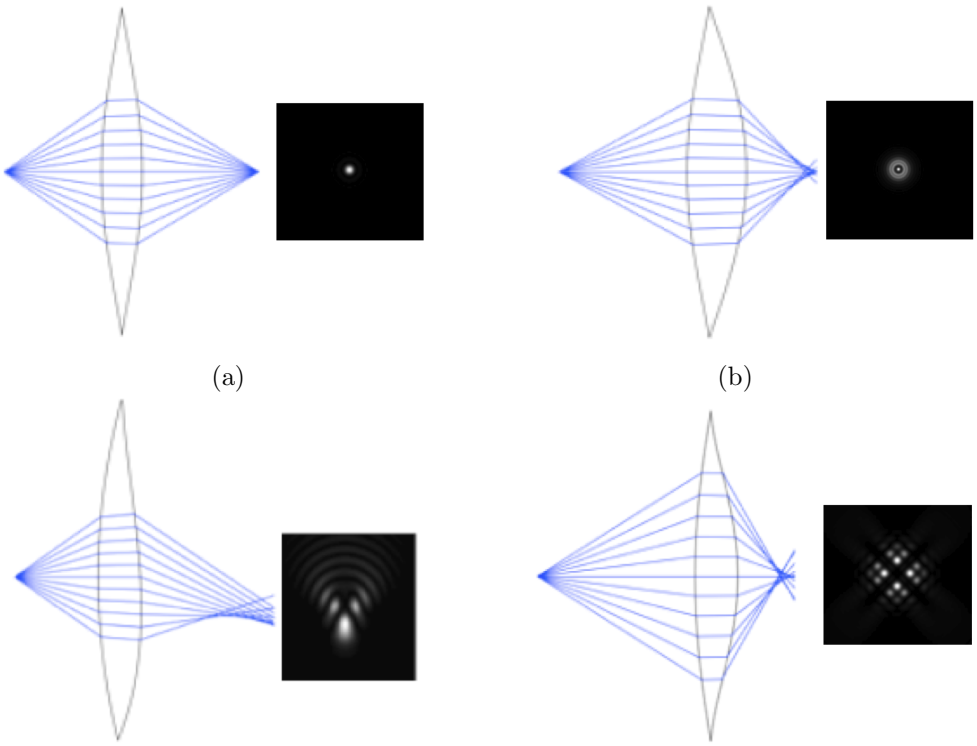

Figure. 2.3: Ray diagrams and corresponding PSFs showing a) no aberration, b) defocus, c) coma and d) secondary astigmatism. Diagrams and images were generated using purpose-written python code that calculated ray angles from first principles. The PSF images were generated using the technique described in Section 4.2, which was also implemented in python.

The wavefront sensor measures the wavefront distortion as a function of position in the pupil plane and the inverse wavefront shape is applied to the wavefront corrector, producing a flat wavefront. This system can be used to capture high resolution, aberration-free images of the retina and to deliver sharp images in to the eye. Liang *et al.* (1997) were the first to test visual performance in combination with an adaptive optics system and they found in an improvement in contrast sensitivity for high spatial frequencies. Other studies have also demonstrated the benefit of correcting higher-order aberrations on contrast sensitivity (Williams *et al.*, 2000; Yoon & Williams, 2002; Guirao, Porter, Williams, & Cox, 2002; Li *et al.*, 2009; Artal, Manzanera, Piers, & Weeber, 2010). By comparison, Charman and Chateau (2003) predict only minor improvements in contrast sensitivity unless the subject has high levels of aberration or is viewing in low luminance conditions where the pupil is large. This is because the eye's wavefront aberrations are small within the central 2 mm diameter of the pupil but increase rapidly with pupil radius and are therefore much larger in low light levels that cause the pupil to dilate (Berny & Slansky, 1969). Dalimier *et al.* (2008) however found that under low light conditions the impact of optical aberrations on contrast acuity is limited by the drop in neural contrast sensitivity. The variability in the eye's aberrations between subjects results in a variability in the improvement of contrast sensitivity in the normal population and in highly aberrated eyes, for example keratoconic eyes that exhibit a conical cornea (Guirao *et al.*, 2002). Improvements to spatial acuity (in terms of a size threshold) have also been shown using stimuli such as tumbling Es (Yoon & Williams, 2002; Yoon *et al.*, 2004), landolt-C tests (Rocha, Vabre, Harms, Chateau, & Krueger, 2007) and letters (Artal *et al.*, 2004; Marcos, Sawides, Gambra, & Dorronsoro, 2008; Rocha, Vabre, Chateau, & Krueger, 2010). Others have investigated the impact of imposing aberrations on vision and found that different types of aberration effect visual acuity to different degrees (Applegate *et al.*, 2002; Applegate, Marsack, Ramos, & Sarver, 2003; Jiménez, Ortiz, Hita, & Soler, 2008; Rouger, Benard, & Legras, 2009) and that the rate of decline in visual performance is related to the amplitude of the aberration (Applegate, Ballentine, Gross, Sarver, & Sarver, 2003; Atchison, Guo, Charman, & Fisher, 2009; Atchison, Guo, & Fisher, 2009).

Ravikumar *et al.* (2010) have investigated the impact of phase changes caused by induced aberrations on letter acuity, finding that visual acuity improved when images were phase-corrected. Recently the effects of aberrations on everyday tasks have been investigated, for example Sawides *et al.* (2010) found that AO correction improves face recognition but not facial expression recognition. Here we have summarised the findings relating aberrations and visual performance, a more comprehensive review of use of adaptive optics for vision science has been published by Roorda (2011).

2.3. Image processing

In Figure 2.3 we saw that aberrations cause distortions in the image. We can characterise the effect of an aberration in an imaging system by examining the changes in the way that it transmits the image.

2.3.1. Point Spread Function (PSF)

The PSF of an optical system is its response to a point source and can be used to analyse image quality. If we assume that the PSF is shift invariant, the image of a complex object is a stationary convolution of the reflectance of that object with the PSF. If the illumination from the source is coherent interference occurs and the PSF is the amplitude of the Fourier transform ($\mathcal{F}$) of the complex pupil function. For incoherent illumination the phase of the light fluctuates randomly and so interference effects are not observed. In this case the PSF ($P(r,\, \theta)$) is the intensity (squared modulus) of the Fourier transform ($\mathcal{F}$) of the complex pupil function:

$$P(r,\, \theta) = |\mathcal{F}(p(r,\, \theta)e^{-i\frac{2\pi}{\lambda}W(r,\, \theta)})|^2, \tag{2.6}$$

where λ is the wavelength of light, $W(r,\, \theta)$ is the wavefront and $p(r,\, \theta)$ is the pupil function. The PSF is then normalised such that the intensity sums to one. In the absence of an aberration the PSF is diffraction limited, with the diffraction pattern being defined by

the size and shape of the aperture. For a circular aperture the diffraction limited angular resolution (θ) of an optical system is given by the Rayleigh criterion:

$$\theta_{dl} = \frac{1.22\lambda}{D} \tag{2.7}$$

where D is the diameter of the aperture. The angular resolution defines the minimum angular separation of two point objects at which the centre of one PSF overlaps with the first minimum of the other PSF. In the presence of aberrations the size and shape of the PSF changes, indicating a reduction in image quality. This can be characterised by metrics such as Strehl ratio and encircled energy.

2.3.2. Optical transfer function (OTF)

The image quality can also be characterised in terms the transmission of different spatial frequencies, as described by the optical transfer function (OTF). The OTF is the Fourier transform of the PSF and is given by

$$OTF(\xi,\eta) = MTF(\xi,\eta)e^{-iPTF(\xi,\eta)}, \tag{2.8}$$

where ξ and η are the spatial frequencies in the x and y directions. $MTF(\xi,\eta)$ is the modulation transfer function (magnitude of the OTF), which represents the contrast with which spatial frequencies are transmitted. $PTF(\xi,\eta)$ is the phase transfer function, which represents the phase shifts with which spatial frequencies are transmitted. In a ideal optical system the MTF has a value of 1 and the PTF has a value of 0 for all spatial frequencies. However, even in the absence of aberrations, diffraction causes a decline in contrast with spatial frequency. The cut-off frequency is the spatial frequency at which the contrast falls to zero and is given by $\frac{D}{\lambda}$. The aperture acts as a low pass filter: increasing its diameter allows higher frequencies to be transmitted by the optical system, improving the angular resolution. In a real optical system, such as the eye, increasing the size of the aperture often

reduces the resolution. This is because the optical quality falls away from the centre of the lens and so aberrations are increasingly introduced. Aberrations do not affect the cut off frequency but they do reduce the modulation depth. If the modulation depth is reduced below the neural contrast threshold then the effective cut off frequency will be reduced. They also cause phase changes, for example defocus produces an OTF with a real part that oscillates between positive and negative values, constituting 180° phase shifts. The contrast of the phase shifted spatial frequencies appears inverted in the image. Other aberrations can cause phase shifts that are not just 0° or 180° and aberrations that produce an asymmetric PSF can cause variations in the PTF. These variations cause the spatial frequencies to be recombined with different relative phases.

2.3.3. Visual strehl

Analysing the OTF is a useful method for calculating the retinal image quality but sampling by the cells in the retina causes an additional weighting to the OTF. By considering the neural weighting on the OTF we can characterise the retinal image quality in terms of its potential implications for visual performance. Although there is no perfect metric for predicting visual performance from a wavefront measurement, Marsack *et al.* (2004) showed that the best metric was the area under the weighted OTF, known as the visual Strehl ratio computed in the frequency domain (VSOTF). This accounts for 81% of the variance in high-contrast logMAR acuity. The VSOTF is defined as

$$VSOTF = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} CSF_N(f_x,f_y)\cdot OTF(f_x,f_y)df_xdf_y}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} CSF_N(f_x,f_y)\cdot OTF_{DL}(f_x,f_y)df_xdf_y}, \tag{2.9}$$

where CSF_N is the neural contrast sensitivity function, OTF is the OTF of the system with an aberration and OTF_{DL} is the OTF of an equivalent diffraction limited system (Thibos, Hong, Bradley, & Applegate, 2004). The contrast sensitivity function of the eye, CSF, is a product of the neural contrast sensitivity function and the contrast transfer function of the

eye's optics, MTF_{eye} (Campbell & Green, 1965). Therefore the neural contrast sensitivity function is defined as

$$CSF_N = \frac{CSF}{MTF_{eye}}. \tag{2.10}$$

Mannos and Sakrison (1974) proposed an empirical model of the contrast sensitivity function

$$CSF = 2.6(0.0192 + 0.114f)e^{-(0.114f)^{1.1}}, \tag{2.11}$$

where f is the spatial frequency.

Equally blurred stimuli of different sizes are not simply scaled versions of each other. Therefore the Zernike coefficients required to produce different sized stimuli with equivalent amounts of blur will be dependent on the size of the stimuli. For a given type and amplitude of aberration, the OTF in retinocentric coordinates (cycles per degree) is the same for any letter size. However, for different letter sizes it is the spatial frequency content in letter-centric coordinates (cycles per letter) that remains unchanged. If the letter is made twice as large, the number of cycles per letter is the same but the number of cycles per degree is halved. In this thesis we assume that the VSOTF predicts visual performance well enough that we can use it as a metric for determining the amplitude of particular aberration that creates an equivalent degradation in the image to that of another aberration. To equate aberration amplitudes across different letter sizes we compare the visual Strehl ratio of those amplitudes by scaling the spatial frequency axis in the OTF so that the resolution in the OTF is the same in each case. The spatial frequency scale in the OTF (s_{OTF}) is given by

$$s_{OTF} = \frac{1}{s_{PSF}N}, \tag{2.12}$$

where N is the number of pixels in the OTF, α is the oversampling, R is the pupil radius and s_{OTF} is the spatial frequency scale in the PSF. Using Equation 4.6,

$$s_{OTF} = \frac{2R\alpha}{\lambda N}. \tag{2.13}$$

Converting from a letter with angular size υ_1 to a letter with angular size υ_2 by a factor of A

$$\upsilon_1 = A\upsilon_2, \tag{2.14}$$

the spatial frequency scale in the OTF changes

$$s_{_{OTF,1}} = \frac{s_{_{OTF,2}}}{A}. \tag{2.15}$$

Using equation 2.12 this gives

$$\frac{2R_1\alpha}{\lambda_1 N_1} = \frac{2R_2\alpha}{\lambda_2 N_2 A}. \tag{2.16}$$

By scaling the OTFs and neural contrast sensitivity functions it is possible to determine the amplitude of aberration for a larger angular size that gives an equivalent amount of blur to the original aberration but over a smaller angular size. Additionally aberrations can be compared over different pupil sizes and wavelengths. Chapter 7 will refer to this calculation when making comparisons between the reading and letter recognition experiments.

2.3.4. Convolution and Cross-correlation

In Section 2.3.1 we referred briefly to the convolution of a PSF with an object to give the image that would be formed by an optical system. In this thesis we will use this to simulate the effect of an aberration on letters and words. Convolution is an integral that represents the amount of overlap between two signals as one is reversed and shifted over the other. For a system in which the PSF is constant across the field of view, convolution between the input signal, *I(x, y)*, and the system's PSF, gives the output signal of the system. It is calculated using the convolution theorem:

$$\begin{aligned} I(x,\ y) \otimes P(x,\ y) &= \int\int I(\mu,\eta) P(x-\mu,\ y-\eta)\, d\mu\, d\eta \\ &= \mathcal{F}^{-1}[\mathcal{F}[I(x,\ y)]\mathcal{F}[P(x,\ y)]], \end{aligned} \tag{2.17}$$

where $\otimes$ indicates convolution.

In this thesis we use cross-correlation to model letter recognition. Cross-correlation in the spatial domain is a measure of the similarity of two different signals as a function of position. It is equivalent to a convolution without reversing the signal:

$$\begin{aligned} I(x,\ y) \star P(x,\ y) &= \int\int I^*(\mu,\eta) P(x+\mu,\ y+\eta)\, d\mu\, d\eta \\ &= \mathcal{F}^{-1}[\mathcal{F}[I^*(-x,-y)]\mathcal{F}[P(x,\ y)]], \\ &= I^*(-x,-y) \otimes P(x,\ y), \end{aligned} \tag{2.18}$$

where * denotes the complex conjugate. Cross-correlation can be used to look for patterns in a signal, for example searching for a match between the template of a letter and a letter candidate. This will be elaborated on in Chapter 4.

2.4. Summary

In this chapter we have described the structure of the eye and the optical components that contribute to ocular aberrations. We have explained how these aberrations are quantified in terms of Zernike polynomials, which form the standard basis set used to describe ocular aberrations. The effect these aberrations have on the retinal image can be described in the spatial domain, via the PSF, and in the frequency domain, via the OTF and VSOTF. Previous studies have shown an improvement in visual performance when higher-order aberrations are corrected, even in the normal population. This increase in performance varies between subjects and is likely to depend on the amplitude of aberrations the subject has initially and their pupil size. In certain conditions it is conceivable that the limitation to visual performance is enforced by neural, rather than optical factors. Additionally the change in performance may be task-specific. The next chapter will cover the neural aspect of vision and visual performance. It will also discuss some of the theory that relates to the visual tasks studied in this thesis.

Chapter 3

The human visual system

3.1. Introduction

The studies of sensation and perception are central to psychology. Sensation refers to the detection of a stimulus, for example the detection of light by the photoreceptors in the retina. Perception refers to the interpretation of the signals from the sensory organs. We sense the light entering our eyes but we perceive the image that the light distribution represents. Perception can be influenced by cognition, which refers to the processes by which we form a coherent representation of the world, recognise its components and remember them. Our perceptions are based on our sensations, so called "bottom-up" effects, but can be shaped by learning and memory, which are examples of "top-down" effects. The important thing to consider is that perception is limited by the sensory information it receives and so can be deceived by erroneous sensory signals.

In the previous chapter we described the structure of the human eye and how the optical components can affect the image quality at the retina. However, the retinal image is encoded in a very complex way by the visual system. It is therefore important to study how the retinal image quality relates to visual performance. In this thesis we have investigated letter recognition and reading. This chapter details some of the theories of visual psychophysics

that relate to these visual tasks and forms a background for concepts referenced in later chapters. Here we will describe the organisation of the visual system, how retinal images are processed and summarise the processes of letter and word recognition. We will also outline some of the theory of the way we move our eyes when we read and how we can use eye movement data to study effects on reading performance.

3.2. Hierarchy of the early visual system

Processing in the human visual system follows a hierarchy that starts with photoreceptors and increases in neural complexity with convergence occurring at every stage. The retina is composed of layers containing different types of cell: photoreceptors, horizontal cells, bipolar cells, amacrine cells and retinal ganglion cells. Melanin is a black pigment required in the retina to prevent reflection within the eye and to restore photo-receptor sensitivity after bleaching. For the melanin to perform this function it must be close to the photoreceptors, which in turn must be connected to the nerve cells that process the signals. This necessitates a front-illuminated design, as shown in Figure 3.1. At the centre of the retina is the fovea, which forms a 0.5 mm shallow pit in which light impinges on the photoreceptors directly in the absence of other retinal cells. The are about 10^4 cone cells of 1 - 4 μm diameter in the fovea that are densely packed (approximately 175 x 10^4 mm^{-2}), giving the highest spatial sampling rate in the retina and hence the greatest acuity.

The two types of photoreceptor cell in the retina are called rods and cones. Rods are sensitive in low luminance (scotopic; $10^{-6} - 10^{-2}$ cd/m^2, peak sensitivity at around 555 nm) conditions and are bleached in day light whereas cones are sensitive in higher luminance (mesopic; $10^{-2} - 1$ cd/m^2 and photopic; $1 - 10^6$ cd/m^2) conditions. There are three types of cone cell that are labelled based on their spectral sensitivity, which is either long wavelength (L-cone; peaking at 575 nm), medium wavelength (M-cone; peaking at 535 nm) or short wavelength (S-cone; peaking at 445 nm) sensitive. Although the three classes have cones

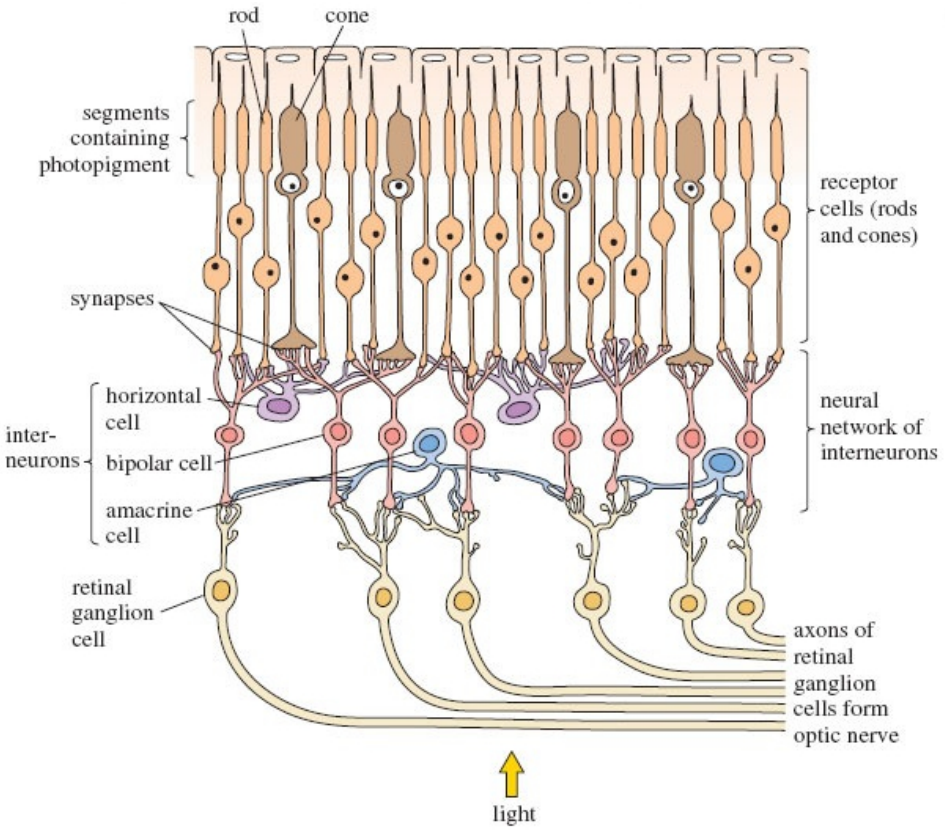

Figure. 3.1: A schematic of the hierarchy of visual processing in the retina showing the light passing through the retinal cells to the photoreceptors. The connections then follow a route from the photoreceptors to the bipolar cells then to the retinal ganglion cells and out via the optic nerve. The horizontal and amacrine cells modulate the retinal receptive field sizes and configurations.(Frisby, 1980)

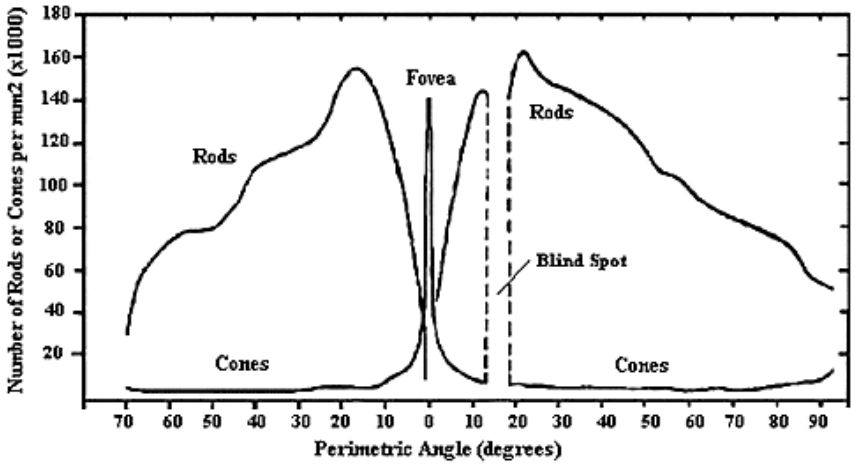

Figure. 3.2: Distribution of rods and cones in a horizontal line across the retina (from Osterberg, 1935).

have different spectral sensitivities, the cone signal does not indicate the wavelength of light by itself. Colour is determined by a comparison of intensity signals from two types of cone. Comparing signals from L- and M- cones mediates red-green discrimination. Blue-yellow discrimination is mediated by a comparison of S-cone signals with a combination of L- and M- cone signals. The distribution of rods and cones is not uniform across the retina, as shown by Figure 3.2 and there are no rods in the centre of the fovea.

Signals from a selection of nearby photoreceptors are combined in a bipolar cell in a

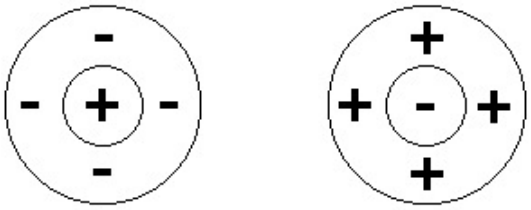

Figure. 3.3: The centre-surround receptive field. In this diagram '−' represents hyperpolarisation caused by incident light and '+' represents depolarisation due to incident light. On the left is a 'on-centre' receptive field with a depolarising centre and on the right is an 'off-centre' receptive field with a hyperpolarising centre.

centre-surround configuration, as shown in Figure 3.3, which in turn feed retinal ganglion

cells. Ganglion cells fire action potentials, which are discrete events that only occur when the cell's potential reaches a certain threshold, and the firing rate represents the strength of the cell's response. The receptive fields of the most common types of ganglion cell (midget and parasol) can be described as either 'on-centre' (illumination in the centre depolarises the cell) or 'off-centre' (illumination in the centre hyperpolarises the cell). This kind of organisation is called spatial antagonism and is important in visual processing as it takes account of the distribution of light and hence the local contrast. An on-centre receptive field under uniform illumination experiences depolarisation in the centre and hyperpolarisation in the surround, causing a net change in potential (from its resting state) of zero and the cell does not fire (in reality it is never actually zero, as explained in section 3.3). However, if the centre is illuminated (so depolarises) but the surround is not illuminated (so does not hyperpolarise) the net change in potential is positive and the cell fires rapidly. This centre-surround configuration is analogous to a kernel in image processing, which is used as method of detecting features. It also equates to image compression since the signals from many photoreceptors are represented in a meaningful way by a single response. The centre of the receptive field follows a direct path from photoreceptors to bipolar cells and the surround follows an indirect path using the lateral connections of the horizontal cells. In the fovea photoreceptors synapse to bipolar cells in a one-to-one manner using only the direct path (the centre), whereas processing in the peripheral retina uses both the direct and indirect paths. This maintains a high rate of spatial sampling in the fovea where acuity is highest. In the peripheral retina receptive field centres are larger indicating a further reduction in sampling rate and hence acuity. Ganglion cell receptive fields tend to be circularly symmetric, though in periphery they may become elongated. The centre-surround configuration of ganglion cell receptive fields is additionally modulated by amacrine cells. There are many different types of ganglion cell, each having their own characteristic wiring arrangement, and believed to be specialised for extracting particular types of visual information (Dacey, Peterson, Robinson, & Gamlin, 2003). Recent work has identified a melanopsin-containing ganglion cell that is intrinsically photosensitive (Berson, Dunn, & Takao, 2002). These cells

depolarise in response to light even in the absence of rod and cone inputs. They are thought to provide the primary photoreceptor input to the suprachiasmatic nucleus of the hypothalamus, which is responsible for circadian rhythm.

The axons of the retinal ganglion cells exit the retina via the optic nerve and cross at the optic chiasm, where the left and right visual fields of each eye are combined and redirected to opposite sides of the brain. Beyond this the optic nerve fibres continue to a few destinations in the brain such as the pretectum, which mediates processes such as pupil reflex, focus and head-related eye movements, and the superior colliculus which is involved in the control of saccadic eye movements (described in Section 3.5.1). Most fibres however continue to the two lateral geniculate nuclei (LGN), one for each side of the visual field. Each nucleus segregates information from each eye into different layers but retains a topographical map of the visual field. The geniculate cells have similar receptive fields to retinal ganglion cells that maintain the segregation of the left and right eye. Fibers from the LGN are combined in the optic radiation and project on to the primary visual cortex. It is in the primary visual cortex where, among other processes, 'bar' and 'edge' detection occurs. Bar and edge receptive fields (see Figure 3.4 and Section 3.4) are derived via selective wiring in which a cortical cell receives organised input from a set of neighbouring LGN cells. This process leads to spatial frequency and spatial phase discrimination, as described in the Sections 3.3 and 3.4.

3.3. Spatial frequency discrimination

The contrast sensitivity function of the human eye was first measured by Westheimer (1960) and by Arnulf and Dupuy (1960). Campbell and Green (1965) compared the optical and neural contributions to the contrast sensitivity function. They concluded that for a pupil diameter less than 2.4 mm the attenuation of high spatial frequencies was mainly due to neural factors unless the aberrations were large. It is possible to see the contrast sensitivity

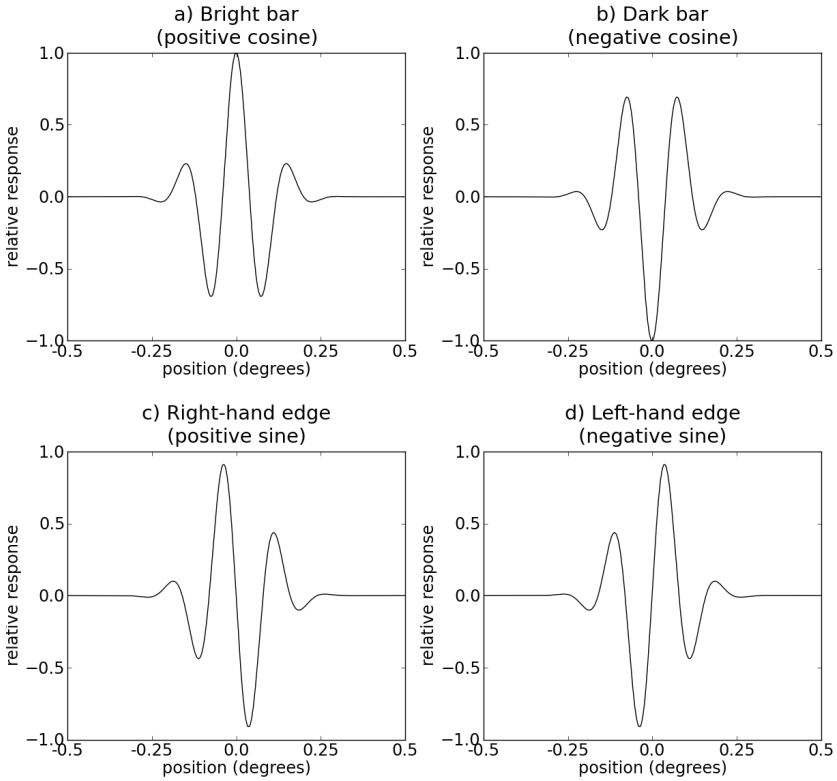

Figure. 3.4: The response functions of the four types of receptive fields in the visual system. These are sensitive to a) a bright bar on a dark background, b) a dark bar on a bright background, c) a bright edge to the left of a dark edge and d) a dark edge to the left of a bright edge. a) and b) constitute the even-symmetric (bar-selective) mechanisms whereas c) and d) are the odd-symmetric (edge-selective) mechanisms. These plots indicate the relative response of a neuron to a point of light as a function of position. Negative regions on these plots indicate an inhibitory effect on the neuron's response whereas the positive regions indicate an excitatory effect. It can also be thought of as the stimulus intensity pattern that would produce the greatest firing rate of the neuron. Diagrams are plotted using equations derived by Marčelja (1980).

function of one's own eye by looking at Figure 3.5, which exhibits a logarithmic variation in contrast against a logarithmic variation in spatial frequency. In this figure a curve can be seen above which the fringes cannot be discerned (see Chapter 2 for the equation describing this curve).

Blakemore and Campbell (1969) were the first to suggest that the human visual system

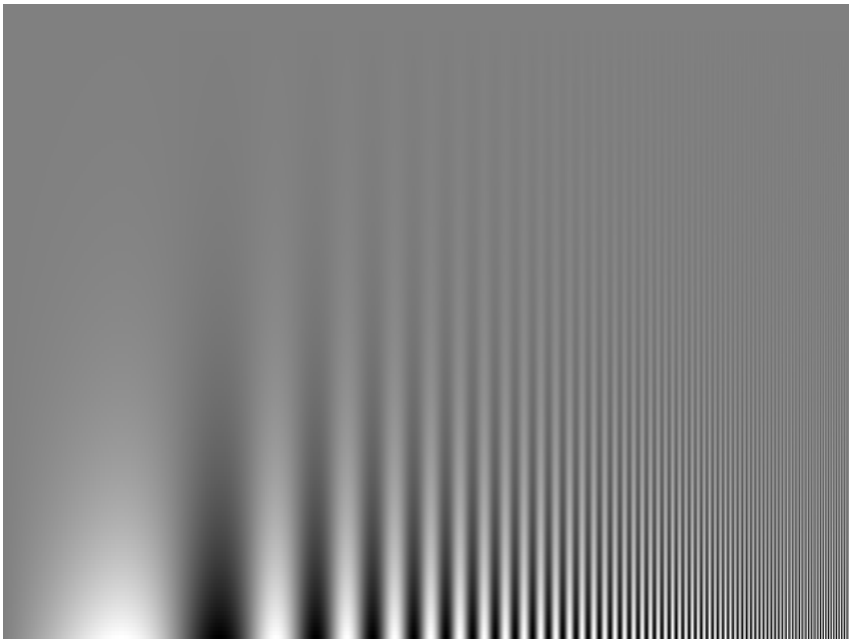

Figure. 3.5: Campbell-Robson contrast sensitivity chart demonstrating contrast sensitivity as a function of spatial frequency (Campbell & Robson, 1964, 1968). Spatial frequency increases logarithmically from left to right and contrast increases logarithmically from top to bottom. A curve can be seen above which the fringes cannot be discerned. Note that the appearance of this chart depends on viewing distance as well as the MTF of the printer or monitor in which it is displayed.

might perform a kind of Fourier analysis. There are a number of separate channels in the visual system that are selectively sensitive to a relatively narrow band of spatial frequencies (Blakemore & Campbell, 1969; Campbell & Robson, 1968; Graham & Nachmias, 1971). The detection of a spatial frequency, using a grating for example, is mediated by the activation of overlapping channels. The channel most closely aligned to the grating frequency will be

activated first and with the highest neuron firing rate, giving a peak response. The response of this channel will determine the contrast threshold at which the spatial frequency can be detected. It has been argued that the contrast sensitivity function is in fact an envelope of a number of these channels. The centre-surround receptive field is well suited to detecting spatial frequencies and it is thought that this occurs in cortical simple cells (i.e., the bar and edge detectors). The spatial distribution of responsivity of both the centre and surround can be described by Gaussian functions. The responsivity of a ganglion cell is the difference of these Gaussians, as shown by Figure 3.6. The centre and surround represent two low-pass filters with different cut-off frequencies and the difference of these two filters is a band-pass filter. In the frequency domain the responsivity is also a difference of Gaussians (the Fourier transform of a Gaussian is a Gaussian), which when expressed on a logarithmic scale leads to the classic contrast sensitivity curve shown in Figure 3.7 and demonstrated by Figure 3.5. The surround is typically less responsive than the centre at very low spatial frequencies and so the difference is never actually zero, even at a spatial frequency of zero (Enroth-Cugell & Robson, 1984).

Letter recognition

Solomon and Pelli (1994) tested the role of frequency channels in letter recognition by superimposing letters over noise characterised by a certain band of spatial frequencies. This work revealed that a noise frequency of 3 cycles per letter masked letter identification with the most efficiency, indicating that letter identification is mediated by a channel sensitive to this frequency. Subjects always use the same one- or two- octave-wide spatial frequency band (in cycles per letter not cycles per degree) for letter identification. When the letter is band-pass filtered the channel frequency scales with the centre frequency of the pass-band (Majaj, Pelli, Kurshan, & Palomares, 2002). This may signify a difference between the spatial frequency channel that mediates sharp text and text viewed with an aberration, which acts to spatially filter the image.

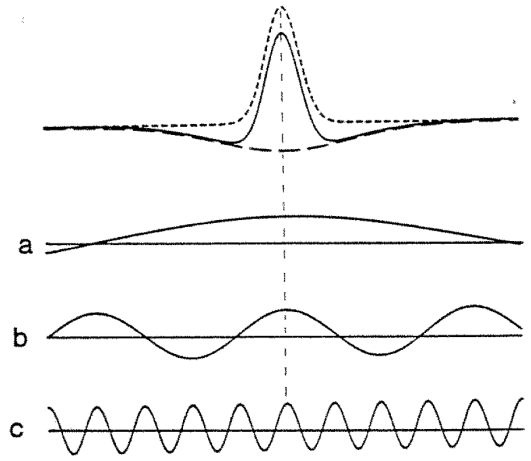

Figure. 3.6: Spatial weighting function of the centre and surround of a receptive field (Enroth-Cugell & Robson, 1984). The short dashes correspond to the centre and the long dashes correspond to the surround. The surround is inverted so that the solid line represents the sum of the two Gaussian components (rather than the difference of Gaussians). Three gratings with different spatial frequencies are shown in a), b) and c).

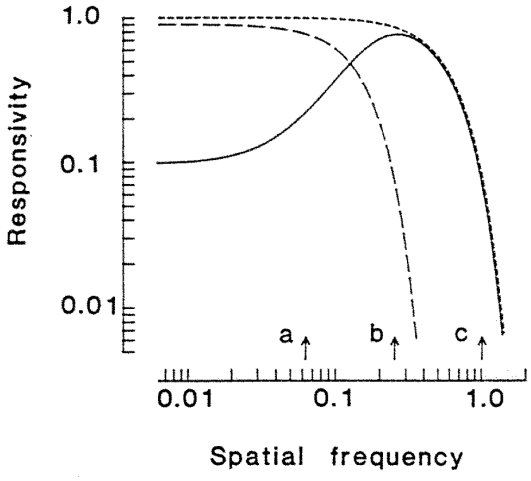

Figure. 3.7: Spatial frequency responsivity function of the centre and surround of a receptive field (Enroth-Cugell & Robson, 1984). As in Figure 3.6 the short dashes correspond to the centre and the long dashes correspond to the surround. The solid line represents the spatial frequency responsivity of a typical ganglion cell which is the difference of the two Gaussian components. The spatial frequency of the three gratings from Figure 3.6 are indicated by a), b) and c).

3.4. Spatial phase discrimination

Spatial phase discrimination uses a broad-band mechanism which is less sensitive than the narrow-band frequency channels (Nachmias & Weber, 1975). The mechanisms for phase encoding are described as even-symmetric (bar-selective) and odd-symmetric (edge-selective) (Field & Nachmias, 1984; Burr, Morrone, & Spinelli, 1989) and are shown in Figure 3.4. The polarity of a bar or edge stimulus can only be determined by the difference in peak response between the two detectors of the channel that is optimised to the target, i.e. the bar or edge channel. In other words, an optimally placed edge detector will detect a bright (positive) bar but the difference in peak response between the positive and negative edge detectors will be negligible and so will reveal no phase information. However, the bar channel will show a large difference in the peak response between the positive and negative bar detectors and so will reveal the phase of the bar. While spatial frequency discrimination is approximately constant across the visual field (Koenderink, Bouman, Bueno de Mesquita, & Slappendel, 1978; Rovamo, Virsu, & Näsänen, 1978), the ability to discriminate spatial phase reduces with eccentricity (Braddick, 1981; Julesz, 1981; Rentschler & Treutwien, 1985). This may be due to a decrease in the sensitivity of the odd-symmetric mechanism (Bennett & Banks, 1987).

3.5. Letter and word perception

It is generally accepted that word recognition is achieved by the parallel processing of letters, rather than letter-by-letter. Reicher (1969) gave compelling evidence for this by showing that letters are recognised faster when they are presented in the context of a word. The most popular model for word recognition, which is based on parallel processing, is the interactive activation model first proposed by McClelland and Rumelhart (1981). In this model the process of word recognition follows a hierarchy of detectors, as demonstrated by

Figure 3.8. The process begins with feature detection, which feeds into letter recognition, which in turn affects word recognition. This process is not one-way as both excitatory and inhibitory messages, which affect the activation rate of their recipients, flow in both bottom-up and top-down directions. Nodes within each of the 3 levels in this diagram represent single units such as a feature, a letter or a word. Connections are only made between nodes in either the same or adjacent levels. Nodes within a level which are connected to each other are called neighbours. Connections within a level are mutually inhibitory since, for example, only one word can occur in one place at a time. If two nodes from adjacent levels are consistent with each other, for example the letter *c* at the beginning of the word *cat*, then their connection is excitatory, otherwise it is inhibitory. Nodes which have a positive activation are said to be active, nodes that are not being fed an input signal are said to be inactive. The resting state of a particular node is determined by its activation frequency over the long term. As an example, words that are common in language (those that have a high lexical frequency) have a higher resting state than less common words (those with a low lexical frequency). Additionally, context information can also influence activation, as demonstrated by Figure 1.2 in Chapter 1, which one reads as "the cat" rather than "the cht" or "tae cat". An example of a network that follows the interactive activation model is given in Figure 3.9. In this example, features such as vertical, horizontal as diagonal lines are detected that lead to increased activation in the *d*, *k* and *r* letter detectors above their resting levels. The other letter nodes are sent inhibitory signals that decrease their activations below their resting level. The active letter detectors increase the activation of the words they are consistent with, for example the letter node *d* sends an excitatory signal to the word node *word* and an inhibitory signal to the nodes *wear*, *work* and *weak*. Once activated the word nodes send feedback to the letter nodes to suppress or reinforce their activation. These activations cascade through the system until the appropriate word has been settled on. If features are detected that are consistent with letters, which are in turn consistent with a word then positive feedback will cause the system to rapidly converge on the appropriate word. Disruptions to feature identification could cause excitatory signals

to letter inappropriate nodes or inhibitory signals to letter nodes that are in fact consistent with the stimulus. This disruption could propagate through the system impairing word identification.

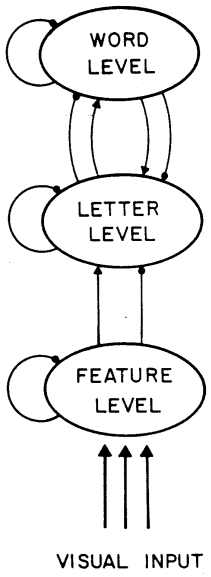

Figure. 3.8: The interactive activation model of word recognition showing a hierarchy of detectors. Excitatory messages (those which increase the activation of their recipients) are indicated with a connecting arrow and inhibitory messages (those which decrease the activation of their recipients) with a connection that terminates with a dot (McClelland & Rumelhart, 1981).

3.5.1. Eye movements and reading

In this thesis we study the effects of aberrations on reading. During reading our eyes make a series of characteristic movements that are driven by the visual processing demands of the task. By analysing these movements we can quantify the differences between types of distortion in terms of their effects on visual and linguistic processing. In this section we summarise the types of eye movement and describe the way that the eyes move during reading.

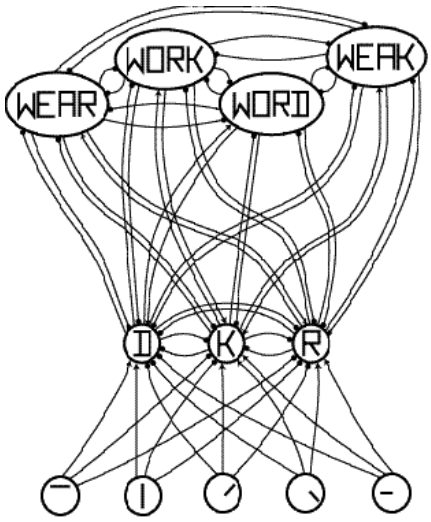

Figure. 3.9: Example of an interactive activation network showing how different feature detectors can activate more than one letter detector which in turn activate multiple word candidates (taken from Bradshaw, 2005).

It is important to consider the effect of higher-order aberrations on a high-level visual task such as reading. Reading differs from clinical acuity tests in two important aspects. In clinical acuity tests letters are presented in isolation and are processed serially. In reading however, words are not processed letter by letter but rather its constituent letters are processed simultaneously so lateral masking and crowding effects between words and letters within words are important. Both of these effects are due to an interference from surrounding letters but while lateral masking refers to a physical overlap of the signals associated with each letter, crowding refers to an effect on the integration of signals in the feature and letter recognition processes that can occur without physical overlap. Secondly, in an acuity test the subject is given an unlimited viewing time to recognise the letter whereas in a reading task the processes of letter and word recognition are time-critical. Disruptions to word recognition do not necessarily completely prevent the subject from reading but do make the task more difficult leading to longer reading times, which can be distressing (see Legge, 2007, for example).

When reading a language such as English, the eyes make a series of ballistic rotations called saccades, each of which moves the point of gaze forwards by approximately 6-9 character spaces (Rayner, 1998). Between saccades the eyes remain quite still, fixating for approximately 200 ms. During fixations visual information is extracted to allow for identification of the word under fixation (Liversedge & Findlay, 2000; Rayner, 1998; Starr & Rayner, 2001). Additionally, since visual input is suppressed during the eye movement, the landing point of the eye movement must be pre-programmed during the preceding fixation. This is achieved via the use of parafoveal information about words to the right of fixation, and in particular their length (McConkie & Rayner, 1976; Rayner & McConkie, 1976). Arguably, the duration of a fixation on a particular word is determined not only by the information extracted foveally while fixating on it, but also by both the extent to which it was parafoveally pre-processed prior to fixation, as well as the ease with which it is identified and interpreted within the context of the sentence or paragraph up to that point. Thus, the fixation duration on a word is affected by the characteristics of that word such as its lexical frequency. It is well known that words that are less common in language take longer to identify (Rayner *et al.*, 2005; Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Rayner, 1998) and this is known as the lexical frequency effect. Characteristics such as word length and the contextual predictability of the word also increase fixation durations, and longer or less predictable words may also require an additional fixation. Another important point to understand is that readers do not process text symmetrically about the point of fixation. It is the case that visual acuity reduces symmetrically with increased horizontal peripheral distance from the fovea. Intuitively, therefore, one might imagine that the same amount of text would be processed to the left, and to the right of the point of fixation. However, this is not the case. The perceptual span (Rayner, 1975) is the number of characters a reader can process during a fixation and is 14-16 to the right of fixation (DenBuurman, Boersma, & Gerrissen, 1981; McConkie & Rayner, 1975; Rayner & Duffy, 1986; Rayner & Bertera, 1979; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981) and 3-4 to the left (McConkie & Rayner, 1976; Rayner, Well, & Pollatsek, 1980; Underwood & McConkie, 1985). The

asymmetry of the perceptual span reflects the importance of attention in reading and how this is centrally associated with what we are processing moment-to-moment during any particular fixation.

It is also known that word spacing affects eye movements and fixation durations and when the spaces are altered in English text it is much harder to read than when they are not (Fisher, 1976; Malt & Seamon, 1978; Morris, Rayner, & Pollatsek, 1990; Pollatsek & Rayner, 1982; Rayner, 1998; Spragins, Lefton, & Fisher, 1976). This could be due to lateral masking effects, or due to the removal of word boundary information that hinders saccadic targeting (saccades are targeted towards the middle of words; e.g., see White & Liversedge, 2004). Physical blurring of letters due to optical imperfections can also affect word spacing such that the spatial extents of words overlap. Disruption to saccade planning, in turn, has consequences for fixation durations, particularly if saccade landing positions are not in the optimal place within a word for its identification. Often when the eyes land in a non-optimal position within a word a corrective eye movement is necessary, resulting in increased processing time on a word.

Considering the average fixation duration (i.e., total sentence reading time in relation to the number of fixations made over a whole sentence) provides a measure of overall reading performance. However it is also possible to consider effects associated with the processing of specific words, and whether such effects occur due to the linguistic characteristics of that word (e.g., its frequency within the language), or because of its visual appearance (e.g., whether it is or is not visually degraded). For example, if two sentences are presented such that the only difference between them is a single high or low lexical frequency target word (matched in length and predictability), the difference in fixation duration on that word may be attributed to the difference in processing time required to identify it. Furthermore, if the text is blurred and the blur changes the appearance of letters in the word such that they look like other letters (i.e., they become more confusable), then visual processing of the word will be more difficult and subsequent word identification will be disrupted (see Section 3.5).

As a consequence, fixation durations will be increasingly affected. This may be particularly problematic if the word is similar to many other words in terms of its constituent letters and has many lexical neighbours (words that differ by a single letter) (Coltheart, Davelaar, Jonasson, & Besner, 1977).

3.6. Summary

In this chapter we described some of the psychophysical processes that underpin the experiments reported in this thesis. We have described the structure of the visual system in terms of its layers of processing cells. The receptive fields of these cells have been discussed in relation to spatial frequency and spatial phase detection. These detection processes form the first stage of letter and word recognition by mediating feature detection. One of the most popular models for word identification, the interactive activation model, has been summarised. The important points to consider are than word recognition depends on accurate letter identification, which in turn is driven by successful feature detection. Disruptions to feature identification could inhibit appropriate letter identification providing the word recognition level with incorrect orthographic information. As described in Section 3.5.1, eye movements during reading are driven by the visual input. If the visual input is degraded, feature detection, letter identification and word recognition are hindered. The lexical frequency effect can be exploited as a way to measure effects on word identification. If the difference between fixation durations on low and high frequency words (the size of the lexical frequency effect) is affected by an aberration we can infer that the aberration affects word recognition, not just the efficient extraction of features. If we accept an interactive activation model for word recognition we can attribute this disruption in word recognition to a reduction in the activation of letter and word nodes, which negatively impacts the feedback mechanisms causing a slower convergence on the appropriate word.

Chapter 4

Simulating ocular aberrations

4.1. Introduction

The effects that monochromatic aberrations have on visual performance can be studied using two approaches. We can investigate either the improvement in visual performance from correcting these aberrations or we can study the impairment to visual performance caused by adding them. In studying the latter we are free to choose which aberrations to investigate for their effects on visual performance. By studying different types of aberration we aim improve the interpretation of ocular wavefront measurements.

Using an adaptive optics system it is possible to induce a particular aberration in isolation. In such a system the subject's ocular aberrations would be corrected in closed loop and the aberration under investigation would be added. Provided that the phase correction was applied in a plane conjugate to the subject's entrance pupil, the image formed at the fovea would be a true representation of the applied monochromatic aberration. We want to study the effects of these aberrations on reading performance, which requires a large amount of eye movement, up to 15 ° with our equipment. There are some technical issues with using an adaptive optics system in this case since large movements of the eye cause the entrance pupil to shift and tilt. The pupil could be stabilised on the wavefront sensor and corrector

using fast scanning mirrors and a control algorithm to interpret the eye movement data. We wanted to allow our subjects to read binocularly, requiring a second adaptive optics path. This could be a separate path with an additional wavefront sensor and corrector or possibly an offset path using the same wavefront sensor and corrector, assuming that they are large enough to incorporate both pupils and the corrector has low actuator coupling. Our subjects read text from a monitor at a relatively close distance so consequently there would be some convergence of the eyes to account for. It has been shown that vergence of the eyes can vary while reading (Liversedge, White, Findlay, & Rayner, 2006). We would require either binocular tracking, which could be achieved with either a second eye tracking path or possibly with software using position measurements from the wavefront sensor(s). Active correction for convergence would require an additional set of scanning mirrors to steer both pupils on to the wavefront sensor(s) and corrector(s).

Instead of designing, building and funding such a complicated system we have chosen to simulate the aberration in the rendering of the stimuli. Since we do not compensate for our subjects' aberrations, the resulting retinal image is not a perfect representation of the applied aberration. We also make no attempt to simulate chromatic aberration. In Section 4.4 we discuss the implications of simulating monochromatic aberrations rather than imposing them optically. In Section 4.2 we detail the method used rendering aberrations. When choosing the types of aberration to test we could have used a database of aberrations typically found in the population. Instead, we wanted to choose aberrations that are likely to have an impact on vision. Using rendered aberrations we studied their effects on the appearance of letters in an attempt to predict visual performance. This model of performance is discussed in Section 4.3. It was not possible to test every type and combination of types of aberration and so we used this prediction to choose three aberrations for investigation. This kept the duration of the psychophysical experiments to a level that was comfortable for our subjects, typically around half an hour to an hour.

4.2. Rendering aberrated stimuli

In our experiments we wanted to simulate the effect of a particular Zernike aberration. Recall from Equation 2.1 that the wavefront of the aberration is given by

$$W(r,\,\theta) = \sum_{n=0}^{N} \sum_{m=-n}^{n} c_n^m Z_n^m(r,\,\theta), \tag{4.1}$$

where c_n^m is the Zernike coefficient corresponding to the Zernike mode $Z_n^m(r,\,\theta)$ and N is the highest radial order to be considered. For programming purposes the wavefront aberration is represented in cartesian coordinates, $W(x,\,y)$ by converting from polar coordinates with

$$x = r\cos\theta \tag{4.2}$$

$$y = r\sin\theta. \tag{4.3}$$

As described in Chapter 2, for incoherent light the PSF created by the particular wavefront is then

$$P(x,y) = |\mathcal{F}(p(x,y)e^{-i\frac{2\pi}{\lambda}W(x,y)})|^2. \tag{4.4}$$

To produce a PSF that samples the wavefront correctly the pupil function and wavefront must be zero-padded to increase the number of frequency samples. The factor by which the number of samples increases is known as the oversampling and this should be at least 2 to maintain Nyquist sampling (Nyquist, 1928; Shannon, 1948). This is because the incoherent PSF is the squared modulus of the coherent PSF and we must pre-compensate for this squaring, which compresses the spatial frequency axis by a factor of 2. Given the PSF of an incoherent optical system, an image produced by the system is simply a convolution of that PSF with the reflectance of a target object. If the field of view under investigation is within the isoplanatic path of the system (that over which the PSF is independent of the field of view), the resulting stimulus image, $I(x,y)$, is given by

$$I(x,y) = O(x,y) \otimes P(x,y) \tag{4.5}$$

where $O(x,y)$ is the reflectance of the target object and $\otimes$ indicates convolution. An example of the process used to generate the stimuli is given in Figure 4.1 and Figure 4.2 shows the effect of rendering the letter e with 0.6 μm of each of the Zernike modes. In our experiments we always consider the field of view to be within the isoplanatic patch (about 1 - 4 °) since the eyes move to keep the fixation point within the fovea. In the reading experiment subjects also used peripheral and parafoveal visual information. The retinal image had off-axis aberrations introduced by the subject's eye which produced a degradation in image quality with visual angle. This degradation would affect peripheral visual information that is important for reading. Since the limit imposed by the photoreceptor density is likely to be the main limitation to peripheral visual acuity it was not necessary to produce a wide-field simulation of the image that would also require real-time, gaze-contingent image rendering.

4.2.1. Calibration of the PSF and target image

Before using a PSF to produce a stimulus image the pixel scales in the PSF and target image must be calibrated to ensure they match. The pixel scale in the PSF, s_{PSF}, is defined as

$$s_{PSF} = \frac{\lambda}{2R\alpha}, \tag{4.6}$$

where R is the radius of the pupil and α is the oversampling factor. If we wish to convolve this PSF with a target letter we need to construct the target such that it has the same pixel scale. The pixel scale of the target object is

$$s_{object} = \frac{\upsilon_{object}}{N_{object}}, \tag{4.7}$$

where υ_{object} is the field of view of the object and N_{object} is the number of pixels in that field of view. Equating Equations 4.6 and 4.7 re-arranging gives the required size of the target in pixels

$$N_{object} = \frac{2R\alpha\upsilon_{object}}{\lambda}. \tag{4.8}$$

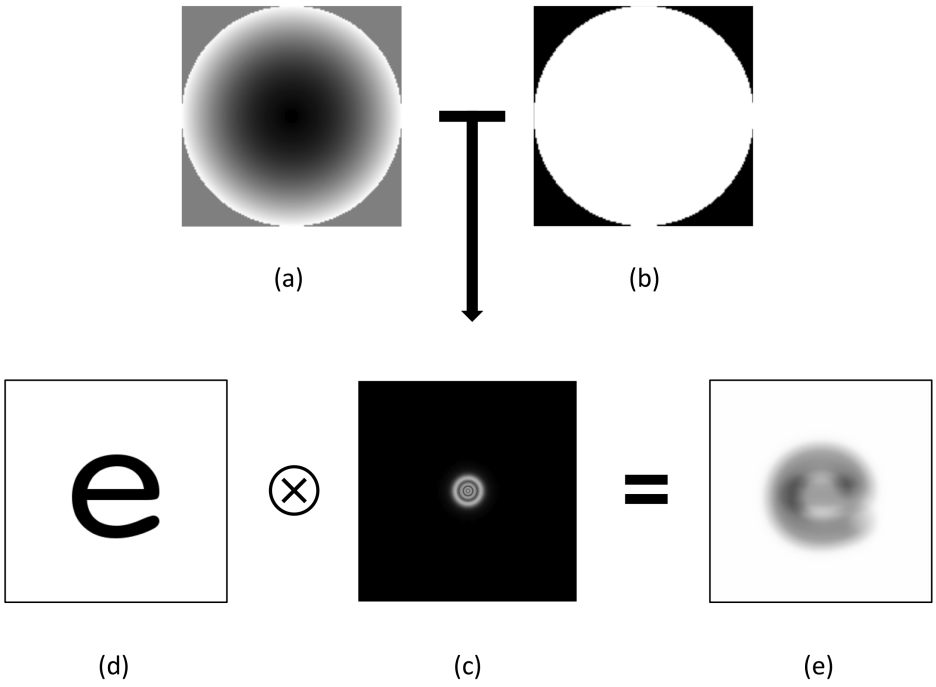

Figure. 4.1: Example of the convolution of the letter e with the PSF of 0.6 μm defocus. This example shows the procedure used for generating the stimuli used in this experiment. The wavefront (a) represents 0.6 μm of defocus and was generated using Equations 2.1 to 2.4 with a radius of 1. This array was padded with zeros so as to double its size in both dimensions from 150 by 150 to 300 by 300 elements. A pupil function (b) was created by setting the elements of an array of 150 by 150 to 1 inside a circle of radius 1 and to zero outside this circle. This was zero-padded to 300 by 300 elements and the incoherent PSF (c) was calculated using Equation 4.4. An array representing the letter e (d) was generated using the Python ImageFont module using Equation 4.8 to determine the correct size so that the pixels scale matched that of the PSF. This target image was convolved with the PSF to give the stimulus array (e).

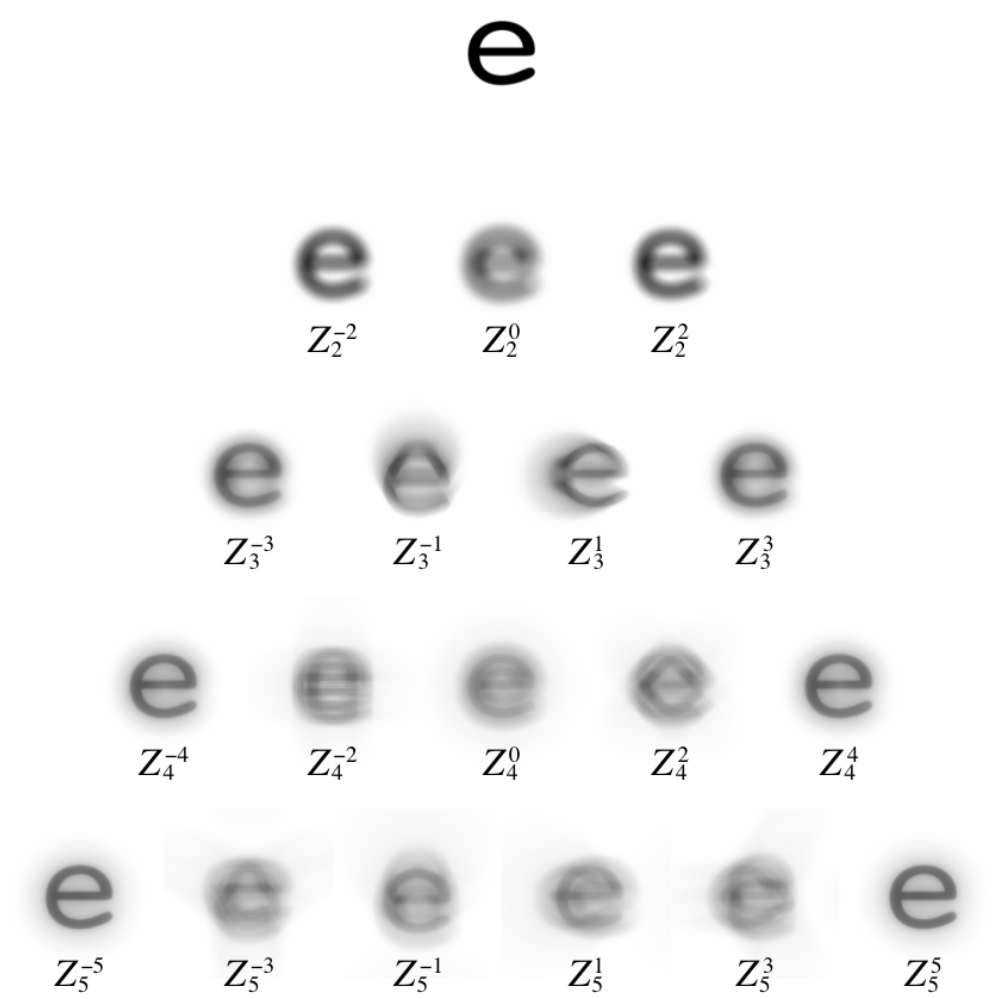

Figure. 4.2: A representation of the Zernike pyramid indicating the effect of 0.6 μm of each mode on a 1° letter *e*. The notation for the Zernike modes is the same as that used in Figure 2.2. The images are scaled such that they are the correct size when held at arms length (48 cm).

A target image of this size will produce the correct stimulus image when convolved with the PSF. Using the pixel size of the monitor and the presentation distance, the stimulus image can then be resized to give the correct field of view for the experiment.

4.3. Choice of aberrations

It is not feasible to test several amplitudes of every Zernike mode so we selected three for investigation by considering their effects on letters and also their clinical relevance. We considered the effect of an aberration on the form of a letter and on the position of the centre of a letter. In this analysis and both experiments we used Courier font as it has small, but significant, advantages to reading acuity and to reading speed in low vision (i.e. vision that is poor despite spectacle or contact lens correction) over other fonts (Mansfield, Legge, & Bane, 1996).

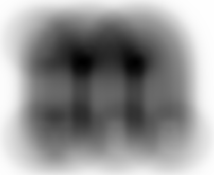

Figure. 4.3: Defocus applied to the letter m such that it appears to contain a letter n, illustrating the effect of a spatial phase change in the image.

4.3.1. Letter confusion

Spatial phase changes caused by aberrations can cause one letter to look like another, as shown in Figure 4.3, where defocus is added to the letter m such that it contains a letter n. Considering this we investigated the potential for letters to be confused by a subject.

We compared letters by performing cross-correlations, giving a measure of similarity based on linear transforms of the stimuli. Examples of this are given in Figures 4.4 and 4.5. For each Zernike mode, an amount of aberration was applied in simulation to each letter of the alphabet. The resulting letters were compared in two ways. First we cross-correlated aberrated letters with other aberrated letters to represent a subject trying to identify letters by looking for differences between them. Secondly we cross-correlated aberrated letters with non-aberrated letters to represent a subject trying to identify letters by using known letter shapes. The maximum value of the cross-correlation was taken to be a measure of 'confusability' as it allows for positional effects. It should be noted that scale effects and other transformations, such as skew, are not accounted for. The values for pairs of letters were entered in a 26-by-26 matrix, which was subsequently normalised such that the values along the diagonal equalled one by multiplying matrix elements by

$$n(x,y) = \frac{1}{\sqrt{c(x,x) \times c(y,y)}}, \tag{4.9}$$

where $n(x,y)$ is the normalisation constant for matrix element (x,y) and $c(x,x)$ and $c(y,y)$ refer to the diagonal matrix elements prior to normalisation for the letters corresponding to x and y. An example of this normalised confusion matrix, calculated for letters with no aberration, is shown in Figure 4.6.

In order to derive a single confusability value we took a weighted mean of the matrix. These weights accounted for the probabilities of letters occurring in language so that, for example, distinguishability of a letter e has a greater effect on the result than that of a letter z. For this we used the letter counts of Jones and Mewhort (2004) which used approximately 183 million English words and counted upper and lowercase letters separately. The mean value of the final matrix was calculated as a metric for the extent to which a particular mode of aberration made letters less distinguishable. These values are represented in the conventional Zernike pyramid in Figures 4.7 and 4.8. Based on these results we chose to investigate secondary astigmatism (Z_4^2) since it produced the greatest confusability factor

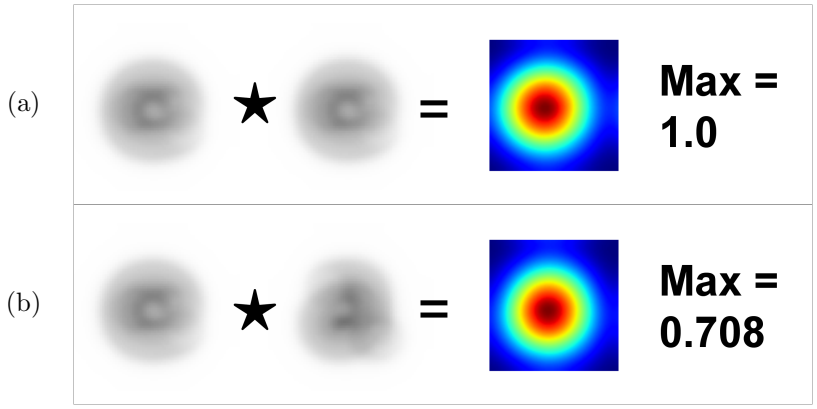

Figure. 4.4: Examples of the cross-correlations ($\star$) performed when calculating the confusion matrix for aberrated vs. aberrated letters. The value on the right indicates the maximum value of the cross correlation, normalised such that the confusion matrix has a value of one along the diagonal. On the left is a letter *e* with 0.4 μm defocus simulated, which could easily be confused with a letter *a*. The letter *e* is cross-correlated with a) itself and b) a letter *a* with 0.4 μm defocus simulated. The peak values produce results for the confusion matrix in Figure 4.7.

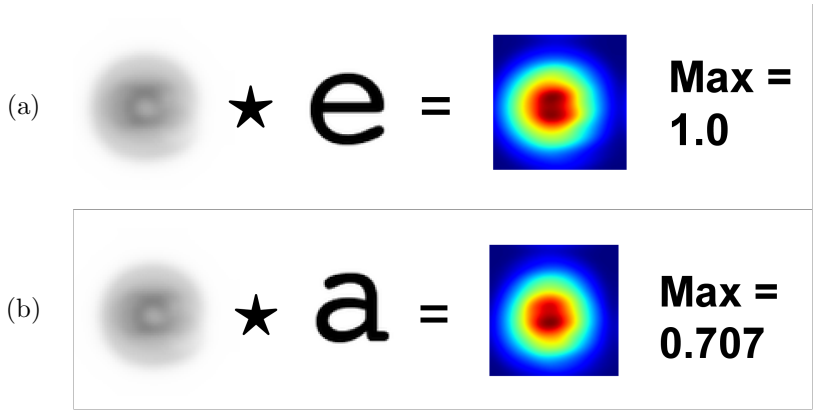

Figure. 4.5: As for Figure 4.4 but comparing aberrated vs. unaberrated letters. The letter *e* is cross-correlated with a) an unchanged letter *e* and b) an unchanged letter *a*. The peak values of the cross-correlation produce results for the confusion matrix in Figure 4.8.

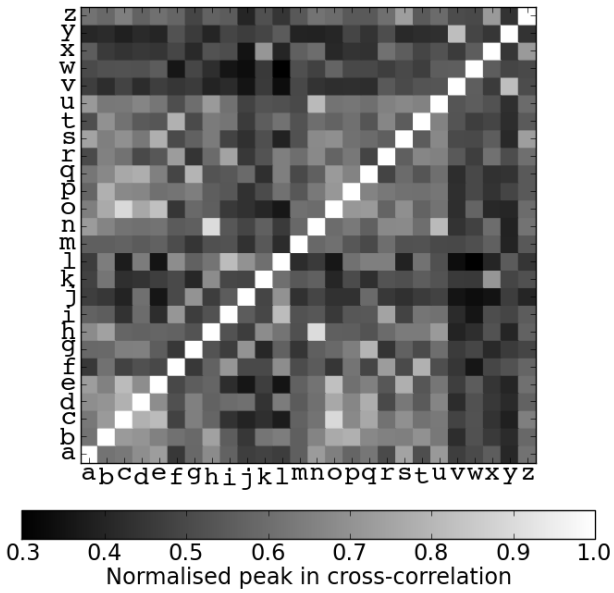

Figure. 4.6: Confusion matrix comparing each letter of the alphabet (using courier font) with every other letter when no aberration is applied. The values in the matrix are the peak value of the cross-correlation of one letter with another and the matrix is normalised such that the values along the diagonal are 1. The scale is such that a value of 1 indicates that the letters are identical.

when comparing aberrated and non-aberrated letters.

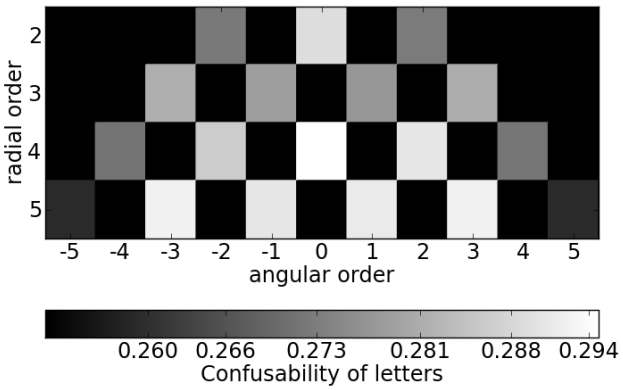

Figure. 4.7: Representation of the Zernike pyramid indicating the confusability when comparing aberrated vs. aberrated letters where a value of 1.0 indicates that letters are completely indistinguishable. This represents a subject attempting to recognise letters by looking for differences between the letters they are presented with. A large value indicates that the aberration makes letters look more similar to each other and therefore more difficult to discriminate.

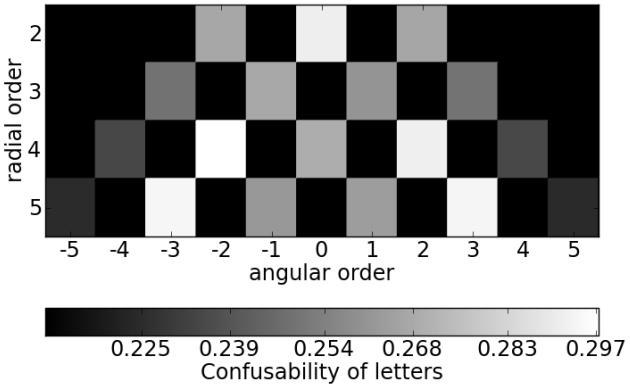

Figure. 4.8: Representation of the Zernike pyramid indicating the confusability when comparing aberrated vs. unaberrated letters where a value of 1.0 indicates that letters are completely indistinguishable. This represents a subject attempting to recognise letters by making comparisons with known letter shapes. A large value indicates that the aberration makes letters look more like the expected appearance of other letters.

4.3.2. Letter position

In addition to changes in the forms of letters, changes in the positions of the centres of letters (and therefore words) were also considered. These are important during reading where lateral masking between letters can interfere with word and letter identification and changes in the word spacings can disrupt saccade targeting. Zernike aberrations that have asymmetric PSFs, such as coma, will cause a shift in the centre of an image. To test the implications of this, each Zernike mode was applied to the letters of the alphabet by simulation and the centres of gravity (the centre of the letter determined by intensity) of resulting images were calculated for the direction parallel to the line of text. The differences between these centres of gravity and those of corresponding letters with no aberration express the changes in letter spacing that occur with each aberration. For each mode the standard deviation of this difference was calculated over all 26 letters. A measure of variance was used instead of an average because if all letters shifted by the same amount the whole word (and sentence) would simply be translated either left or right. It is the differences in the letter spacings that is important for masking effects. These results are represented in Figure 4.9. Coma (Z_3^1) caused the letter centre to vary the most and although the standard deviation is relatively small (1.3% of the average width of a letter) changes in the centres of letters of up to 12.3% the width of a letter could occur in text. Secondary trefoil and secondary astigmatism also cause some variability in the centre of gravity but only about half as much as coma.

In summary, the Zernike modes we have chosen to investigate are coma (Z_3^1) and secondary astigmatism (Z_4^2). Defocus was also chosen for a comparison with a low order aberration and because it relates directly to dioptric blur.

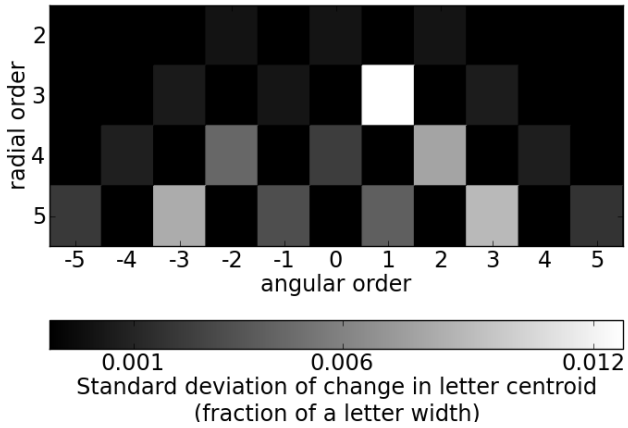

Figure. 4.9: Representation of the Zernike pyramid indicating the standard deviation of the shift (along the horizontal axis) in the centres of gravity of letters (from the no aberration condition) caused by 0.4μm rms of each mode. Shifts are expressed as a fraction of the average width of a letter. The Zernike mode that causes the most variation in letter position in a direction parallel to the text (and therefore the greatest variation in the spacing between letters) is coma (Z_3^1) with a standard deviation of 1.3% of the width of a letter. Secondary trefoil (4.2% for Z_5^{-3} and 3.1% for $Z\%_5^3$) and secondary astigmatism (4.5% for Z_4^{-2} and 7.8% for Z_4^2) cause smaller variability in the centres of letters.

4.3.3. Clinical considerations

From a clinical perspective, secondary astigmatism is thought to contribute to halo phenomena which are often experienced under night vision after successful laser in situ Keratomileusis (LASIK) surgery (Villa, Gutiérrez, Jiménez, & González-Méijome, 2007). Coma is interesting because its amplitude is known to increase with age, which is believed to be due to the loss of compensatory effects between the corneal surfaces (Guirao, Redondo, & Artal, 2000; Lu *et al.*, 2008). It is also prevalent in subjects with keratoconus, for which it is even being considered as a diagnostic quantity (Gobbe & Guillon, 2005).

In summary, the Zernike modes we have chosen to investigate are coma (Z_3^1) and secondary astigmatism (Z_4^2). Defocus was also chosen for a comparison with a low order aberration and because it relates directly to dioptric blur.

4.4. Implications of applying the aberration in the rendering

Since we impose our aberrations in the rendering of the stimuli, we must consider the differences in the retinal image produced compared to applying the aberration optically. When viewing the simulated aberration, the subject's eye introduces additional changes in the retinal image due to monochromatic and chromatic aberrations. We must also consider changes in the eye's aberrations under accommodation. These effects are described in the following sections.

4.4.1. Contribution of subject's aberrations to retinal image

Subjects' individual aberrations produce a different image at the retina to the image produced by a system with the aberration alone. The effect of an optical system on an object described as

$$\mathcal{F}(I(x,y)_{simulated}) = \mathcal{F}(O(x,y))OTF_{system}. \tag{4.10}$$

If that image is then transformed by a second optical system, in this case the subject's eye, the new image formed at the retina is described by

$$\mathcal{F}(I(x,y)_{retina}) = \mathcal{F}(I(x,y)_{simulated})OFT_{_{eye}}, \tag{4.11}$$

where

$$OFT_{_{eye}} = OTF_{pupil}OTF_{eye's\ aberrations}. \tag{4.12}$$

In order to minimise the effect of the subject's eye, the magnitude of its OTF should be close to one for all spatial frequencies. First we consider a diffraction-limited eye that is free of aberrations. The pupil of the eye has an approximately linearly decreasing MTF with a cut-off frequency that increases with increasing diameter. It is therefore desirable, in this case, to have as large a pupil as possible. However, in a real eye the amplitude of aberration increases with increasing pupil diameter. We therefore want as small a pupil as possible to minimise the effect of the subject's eye on the retinal image. To check that this effect was as small as possible, we measured our subject's aberrations using a Zywave$^{\mathrm{TM}}$aberrometer and accounted for their pupil size during the experiment. Different subjects participated in each of the experiments presented in this thesis therefore aberration data are quoted separately in Chapters 7 and 6. Although simulating an aberration produces a retinal image that includes the effects of a subject's own monochromatic aberrations, previous work has shown that the visual system adapts neurally to the eye's particular aberrations. Artal *et al.* (2004) showed that when asked to choose the sharpest image subjects selected those viewed with their own aberrations. Even though our subjects have different ocular aberrations, they are all using the same baseline: the aberrations that they are used to.

4.4.2. Effects due to accommodative error

It is known that monochromatic aberrations change with accommodative behaviour and so we must consider the effect of a subject's accommodative state on the retinal image. Ac-

commodative behaviour changes with visual acuity (see Heath, 1956, for example) and so we expect our subjects' accommodative state to vary with different aberrations. This is because both blur and angular size are strong accommodation cues (Campbell & Westheimer, 1959). The accommodative response of our subjects may therefore be incorrect for the stimulus distance and this would increase the amount of defocus and other higher-order aberrations, particularly spherical aberration (Ninomiya *et al.*, 2002; Cheng *et al.*, 2004; López-Gil *et al.*, 2008). The stimuli used in the reading experiment were less than 1 D of equivalent defocus and they were presented at 75 cm. From the work of Heath (1956) we have estimated the maximum error in accommodation to be 0.25 D, which is equivalent to 0.1 μm defocus over a 3.5 mm pupil. The stimuli used in the letter recognition experiment were less than 4 D of equivalent defocus and they were presented at 170 cm. We have estimated the error in accommodation response to be less than 0.1 D, which is equivalent to 0.02 μm defocus over a 2.5 mm pupil. We expect the actual effect of accommodative error to be much smaller than we have stated here as there were other visual cues available to drive appropriate accommodation. Specifically, the edges of the monitor were visible and in the case of the letter recognition experiment subjects also saw a high contrast box that surrounded the stimulus, which was preceded by a fixation cross. In the reading experiment stimuli were preceded a fixation cross and followed by a comprehension question. These high contrast, high spatial frequency objects should aid the accommodation response. Our subjects viewed the text binocularly and so had vergence and disparity cues available to them. At larger viewing distances accommodation changes have a smaller influence and since pupil size was small the depth of field is large, minimising the effect of changes in accommodation. We also expect polychromatic images to provide chromatic cues to accommodation, as described in the following section.

It is known that as the eye accommodates to nearer objects there are three responses: The curvature of the lens increases, the oculomotor muscles increase the convergence of the eyes and the pupils constrict. This is known as the accommodative triad (Myers & Stark,

1990). During the reading task pupil size was monitored and so it is possible to infer the accommodative state of our subjects. Pupil size also varies with mean luminance, however the mean luminance should not change between aberration conditions. Each condition used the same sentences (in a random order) and the aberration itself does not affect mean luminance, only its distribution. Figure 4.10 shows the sizes of subjects' pupils and a 3 (types) x 3 (amplitudes) two-way repeated measures analysis of variance (ANOVA, see Appendix 1 for a summary) showed that these did not vary significantly between the different types ($p = 0.546$, $MSE = 0.001$, $F(2,0143.6) = 0.6$) or amplitudes ($p = 0.275$, $MSE = 0.005$, $F(2.0,143.6) = 1.3$) of aberration used in the reading experiment and there was no significant interaction ($p = 0.504$, $MSE = 0.002$, $F(2.0,143.6) = 0.8$). Although there appears to be an increase in pupil radius with increasing amplitudes of defocus, this was not significant ($p = 0.126$, $MSE = 0.007$, $F(2.0,36.0) = 2.2$). Subjects' pupil sizes were larger when reading aberrated text than in the control condition suggesting that they may have had a decreased accommodative response. This is consistent with the findings of Heath (1956) that blurred stimuli cause subjects to accommodate slightly further away than the correct distance, possibly because there is no benefit from using more accommodative effort.

The effect that diffraction, higher-order aberrations and accommodative error have on the simulated image is demonstrated in Figure 4.11, which shows minimal changes from the original stimulus. These effects are demonstrated using measurements relating to the reading experiment since those are likely to have a larger effect on the retinal image than those relating to the letter recognition experiment.

4.4.3. Chromatic effects

Longitudinal (axial) chromatic aberration is caused by a wavelength-dependent variation in focal distance. Longer wavelengths are brought to focus at a further axial distance than shorter ones, causing a different PSF. Therefore the image produced in simulating that aberration depends on the wavelength spectrum of the light source. For a polychromatic

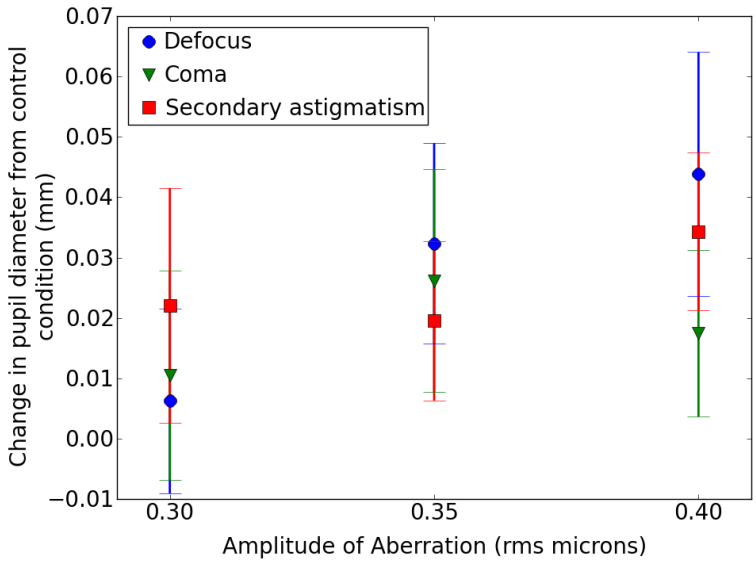

Figure. 4.10: Change in pupil size when reading aberrated rather than clear text averaged over 19 subjects. Pupil size did not vary significantly with the type ($p = 0.546$, MSE = 0.001, $F(2,.0143.6) = 0.616$) or amplitude ($p = 0.275$, MSE = 0.005, $F(2.0,143.6) = 1.339$) of aberration and there was no significant interaction ($p = 0.504$, MSE = 0.002, $F(2.0,143.6) = 0.841$). On average pupils were larger when reading aberrated text than clear text, perhaps suggesting that subjects accommodated to a distance slightly behind the monitor, which is in agreement with the work of Heath (1956).

(a) (b) (c) (d) (e)

Figure. 4.11: The letter *e* presented as a) the original image, b) the image simulated with 0.4 μm defocus, c) the simulated image with diffraction effects from the subject's pupil, d) the simulated aberration with diffraction effects and the subject's higher-order aberrations (averaged over all subjects the amplitude of higher-order aberrations was 0.018μm) and e) the final image at the subject's retina including diffraction, higher-order aberrations and accommodative error (equivalent to 0.1 μm defocus). The effect of an average subject's pupil, aberrations and expected accommodative error on the simulated image are minimal.

source, a monochromatic PSF can be calculated for each wavelength in the spectrum and then the polychromatic PSF is calculated as the linear sum of these PSFs weighted by the luminance at the corresponding wavelength (Marcos, Burns, Moreno-Barriusop, & Navarro, 1999). However, convolving the polychromatic PSF with the stimulus object only produces a correct polychromatic image for spectrally homogeneous sources (Barnden, 1974). These are sources which are spatially uniform in terms of their spectrum, except for a linear intensity scaling. The red, green and blue phosphors of a colour display are individually spectrally homogeneous but are generally inhomogeneous in combination with each other (i.e. there may be colour changes across the image). It is therefore necessary to create a polychromatic PSF for each phosphor and produce a three-colour bitmap of the resulting image (Ravikumar, Thibos, & Bradley, 2008). The image produced, shown in Figure 4.12, has colour changes with the white background tending towards a red hue. To avoid introducing colour effects that may affect reading performance, particularly for longer wavelengths (Matthews, 1987), we chose to simulate the monochromatic aberration only. We chose the wavelength for PSF calculation by considering the spectrum of the monitor, shown in Figure 4.13. The peak wavelength in the spectrum is 627 nm which accounts for 2.5% of the total spectral radiance.

Transverse (lateral) chromatic aberration is caused by the dispersion of off-axis rays producing wavelength-dependent shifts and magnifications of the retinal image. These effects do not occur on-axis and they increase with increasing retinal eccentricity. The transverse chromatic aberration experienced by our subjects was different from that with an optical distortion due to the different light distribution at the eye's optical surfaces. Transverse chromatic aberration is thought to have only a small affect in the near periphery and should not impair peripheral colour vision significantly (Ogboso & Bedell, 1987).

All chromatic aberration could be eliminated by using monochromatic stimuli but we decided against this in order to simulate natural viewing conditions. In addition, longitudinal chromatic aberration changes with accommodation and there is evidence that this may

Figure. 4.12: Images (R,G and B) generated with a spectrally weighted polychromatic PSF computed for each phosphor. The linear sum of these (W) gives the correct retinal image accounting for longitudinal chromatic aberration. The faded appearance of B is caused by the luminance scaling of the image (taken from Ravikumar *et al.*, 2008).

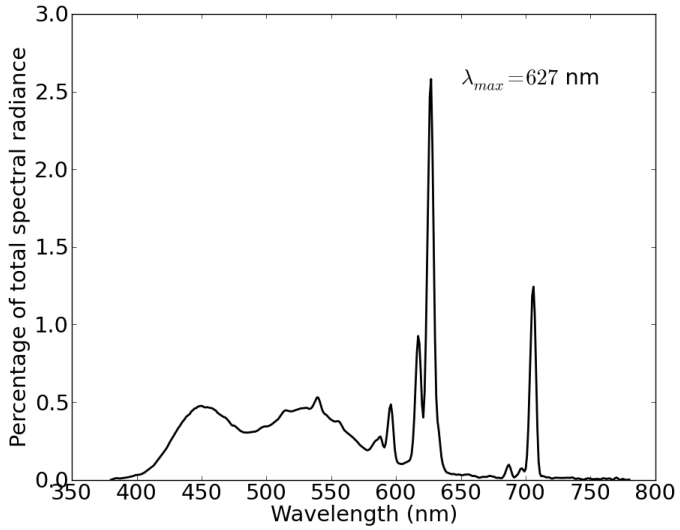

Figure. 4.13: Wavelength spectrum of the monitor when presenting white pixels. The peak wavelength is 627 nm.

aid the accommodation response in some subjects (see Aggarwala, Nowbotsing, & Kruger, 1995; Kruger & Pola, 1986; Kruger, Mathews, Aggarwala, & Sanchez, 1993, for example). Using broadband stimuli, as we did in our experiments, may have been beneficial as our stimuli required our subjects to accommodate correctly.

4.5. Conclusions

This chapter described the process of applying an aberration in the rendering of the stimuli. Implications this has for the results of psychophysical experiments that use such stimuli were also discussed. We concluded that the effect of applying the aberration in the rendering of the stimuli, rather than optically, was minimal. We performed a cross-correlation analysis using rendered aberrations to predict which aberrations would have the greatest effect on letter recognition and investigated their impact on the letter spacings. From these results we chose to investigate the effects of defocus (Z_2^0), coma (Z_3^1) and secondary astigmatism (Z_4^2). Clinically, secondary astigmatism is thought to be involved in halo phenomena and coma is associated with keratoconus. The following two chapters describe psychophysical experiments conducted with these stimuli. The first studies the effect of these aberrations on reading performance and the second examines their effect on letter recognition.

Chapter 5

Reliability of ocular wavefront measurements

5.1. Introduction

The motivation behind this thesis is to understand how different types of aberration affect visual performance. This will lead to a better interpretation of ocular wavefront measurements and an insight into the potential benefits of correcting aberrations. Before we can consider compensating for these aberrations with a static correcting device we need to be sure that we can measure them reliably and that they do not change significantly over time. If the measurements of certain Zernike modes have a high repeatability errors and those modes significantly impact vision then these errors will lead to an unsatisfactory correction. This is also important for the psychophysical experiments described in Chapters 7 and 6 since we rely on these measurements to account for the effect of subject's own aberrations, as described in Chapter 4. It is therefore essential that we not only understand the repeatability errors associated with Shack-Hartmann wavefront sensors in general but also to characterise the errors associated with the specific device we have used. We need to know that our ocular wavefront measurements are accurate and repeatable. In this chapter we

study the repeatability errors associated with the Zywave (I) aberrometer in terms of the standard deviation (SD, absolute repeatability errors) and coefficient of variation ($\frac{SD}{mean}$, relative repeatability error) of repeated measurements.

In this chapter we present a series of experiments aimed at understanding the repeatability errors in wavefront measurements made with a Zywave aberrometer. The device is described in Section 5.2 and previous work on measuring its repeatability is detailed in Section 5.3. In Section 5.4 we discuss the repeatability errors measured with a model eye. These results indicate the system-induced repeatability errors and the operator-induced errors. We also quantify the errors associated with changing the pupil radius and focus error (analogous to a myopic or hyperopic shift). In Section 5.5 we discuss the subject-induced repeatability errors and how subject's wavefront aberrations change over time.

5.2. The Zywave aberrometer

The Bausch and Lomb Zywave aberrometer is an ocular wavefront sensing device based on a Shack-Hartmann sensor. It projects a 785 nm laser beam on to the subject's retina that acts as a guide star. This light is scattered back out of the eye and through a lenslet array producing a Shack-Hartmann spot pattern on the CCD. Correct conjugation of the lenslet array to the eye's pupil is achieved by the operator centering and focusing the iris on the CCD. Before each measurement the subject is asked to blink to ensure that their tear film is always in the same state (most importantly ensuring that it has not broken down) and also to discourage them from blinking during the measurement. We noticed that pupil dilation took about a second to settle after a blink so the measurement was initiated after this interval. The measurement is performed in two stages during each of which the subject must keep their eyes open and still. Firstly a Badal lens system (one that changes focus without changing magnification) is used to remove the focus from the wavefront and secondly the Shack-Hartmann spots are recorded. The spot displacements

from their reference positions are used to determine the local wavefront slopes in each subaperture. Zernike polynomials are fit to the slope data and Zernike coefficients up to fifth order are calculated. The wavefront measurement takes 15 ms during which the laser is flashed five times to get five repeated measurements. The Zywave software rejects the two least repeatable measurements and calculates an average of the remaining three. A schematic of this principle created in Zemax is shown in Figure 5.1.

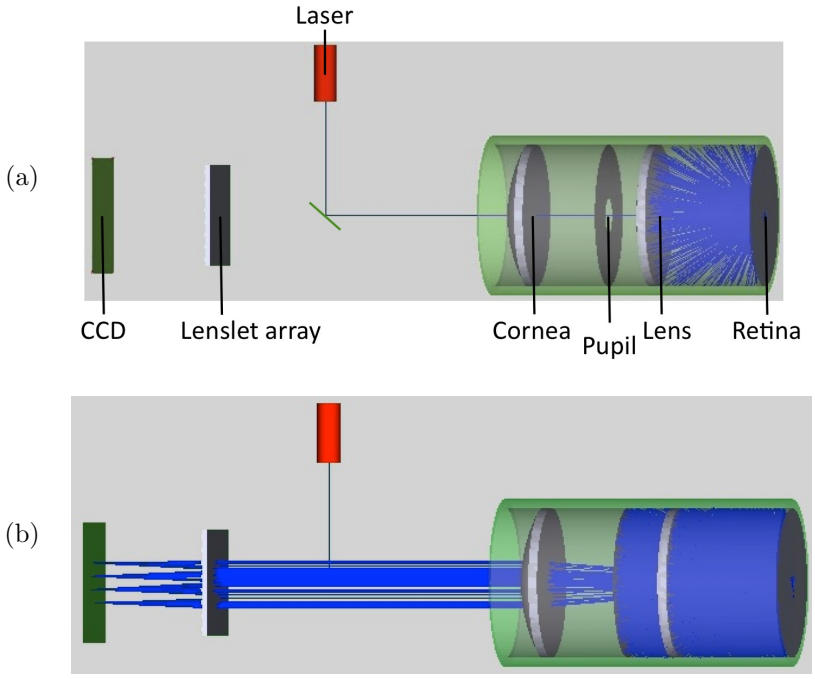

Figure. 5.1: Schematic showing the Zywave aberrometer measurment principle, demonstrated using a model eye (described in section 5.4.1) in Zemax. Panel (a) shows a narrow laser beam incident on the retina that scatters back towards the front of the eye. Panel (b) shows the return light which passes through the wavefront sensor. The Badal lens system used to ensure the light is collimated on the wavefront sensor is not shown.

5.3. Previous studies on the repeatability of Zywave measurements

Previous studies of the Zywave aberrometer have shown that it can provide a repeatable estimate of refractive error. Bullimore *et al.* (2003) and Dobos *et al.* (2009) calculated the

95% limits of agreement (95% *LoA* = *mean difference* + 1.96 *SD of differences*) of two sets of Zywave measurements made on the same set of subjects and found this to be -0.4 to 0.5 D for sphere (focal power). Hament *et al.* (2002) calculated the repeatability coefficient (1.96 *SD of differences*) to be ±0.29 D for sphere, ±0.29 D for cylinder (astigmatism) and ±0.25 D for spherical equivalent (*effective focal power* = *spherical* + $\frac{1}{2}$*cylinder*). Mirshahi *et al.* (2003) calculated the coefficient of variation ($\frac{SD}{|mean|}$) as 0.15 D for sphere and 0.16 D for cylinder.

When measuring higher-order aberrations with the Zywave Mirshahi *et al.* (2003) found that the standard deviation of the total rms amplitude decreased with increasing radial order (0.20 μm for second order, 0.10 μm for third order, 0.06 μm for fourth order and 0.04μm for fifth order). However, the amplitude of aberrations also decreased with increasing radial order. They showed that the coefficient of variation (relative repeatability error) actually increased with increasing radial order (2.2% for secondary astigmatism, 18.4% for third order, 15.8% for fourth order and 29.0% for fifth order). Dobos *et al.* (2009) found that the 95% LoA decreased with increasing radial order. As an alternative measure that is not directly affected by the small amplitudes of higher-order aberrations they compared the standard deviation between sessions to the standard deviation between subjects. They showed that for low order aberrations this ratio was less than 0.1, for third order aberrations it was about 0.2, for fourth order (excluding spherical aberration) it was about 0.4 to 0.9, and for fifth order it was around 0.6 to 1.2. They concluded that for fourth and fifth order aberration measurements the variation between sessions was similar to the variation between subjects and therefore they could not be considered repeatable. Burakgazi *et al.* (2006) and Liang *et al.* (2005) tested the Zywave against other aberrometer devices (NIDEK OPDscan, VISX CustomVue wavefront analyzer, VISX WaveScan and Alcon LADARWave) and found that it had a comparable repeatability error although it gave a significantly different absolute measurement. Burakgazi *et al.* (2006) also concluded that all three aberrometers (OPD-scan CustomVue and Zywave) tested were reliable but this was based on the absolute repeatability error. In this chapter we discuss repeatability measurements although there is

scope to extend this to accuracy measurements by inserting phase plates into the model eye.

5.3.1. Sources of repeatability errors

Operator-induced measurement variance

Mirshahi *et al.* (2003) compared measurements on human subjects with measurements on a test device and found a similar degree of variation in both sets of measurements. They concluded that operator dependent errors were responsible for the variability of their results. These errors could be due to z-axis adjustments, which change the plane that the wavefront sensor is conjugate to, or to x and y axis adjustments made if the device is repositioned between measurements. Cervino *et al.* (2007) tested the effect of operator induced errors by calculating the repeatability of measurements made with a model eye. Using a Carl Zeiss WASCA analyser they measured the model eye's aberrations with and without realignment between measurements. They showed that realignment increased the standard deviation of repeated measurements by a factor of 3. Cheng *et al.* (2003) tested the repeatability of a Wavefront Sciences COAS and found a factor of 2 increase in the standard deviation of repeated measurements with realignment. They also found the device to be tolerant to small axial and lateral misalignments in model eyes, however this may not necessarily be true for highly aberrated human eyes.

Subject-induced measurement variance

In addition to misalignments caused by the operator the same types of misalignment could be caused by the subject moving during the measurement. In this case measurement variation may be induced by x, y and z displacements of the subject's eye (caused by head movements) but also by rotation of the eye (caused by changes in fixation position).

Intrinsic measurement variance

In addition to variance induced by the subject there are other sources of variance caused by the subject that they have no control over. Changes in pupil size can cause significant variance in the measurements. Other sources of variance could be changes in the tear film (Koh *et al.*, 2002; Dubra *et al.*, 2004) and microfluctuations in the accommodative state of the eye (Artal, Fernández, & Manzanera, 2002). Han *et al.* (2007) also suggest that ocular aberration measurements can be affected by eye lash position and by eyelid position since squinting can distort the corneal surface. However these influences should be obvious to the operator when performing the measurement. It is also possible that variations in the wavefront may occur as a result of scattering from the retina, an affect that is inherent in double pass measurements in the eye (Diaz-Santana & Dainty, 2001).

System-induced measurement variance

Once operator-induced and subject-induced errors have been accounted for, the remaining errors must be due to the instrument itself. These repeatability errors are likely to be caused by the centroid estimation which ultimately affects the gradient estimation and therefore the wavefront fitting. The centroid estimation may be affected by the signal to noise ratio (SNR) of the detector, its readout noise and pixel properties. Any noise in the centroid estimation will be amplified or damped by the wavefront fitting. The quality of the Shack-Hartmann spots may also vary since they are formed by scattered rather than reflected light and this will also affect the centroid estimation.

5.4. Measurements with a model eye

To measure the system-induced variance repeated measurements were taken on a model eye, described in the following section. This model was also used to quantify some of the operator-induced and subject-induced variance by taking repeated measures at fixed pupil

radii, fixed focus errors and with different operators. The wavefronts and corresponding Zernike coefficients showing some of these variations are given in Figures 5.2, 5.4, 5.5 and 5.3.

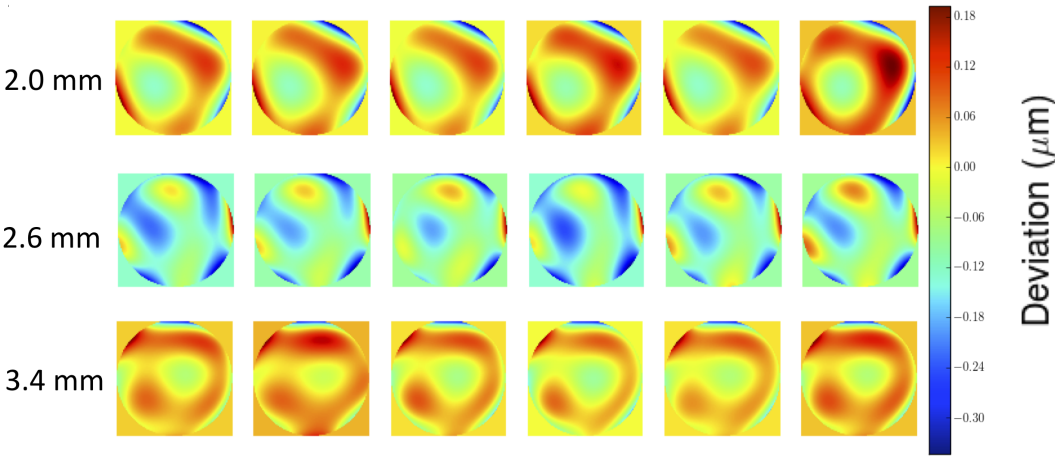

Figure. 5.2: Six repeated wavefront measurements showing variation over different pupil radii with a model eye. Only high-orders are shown.

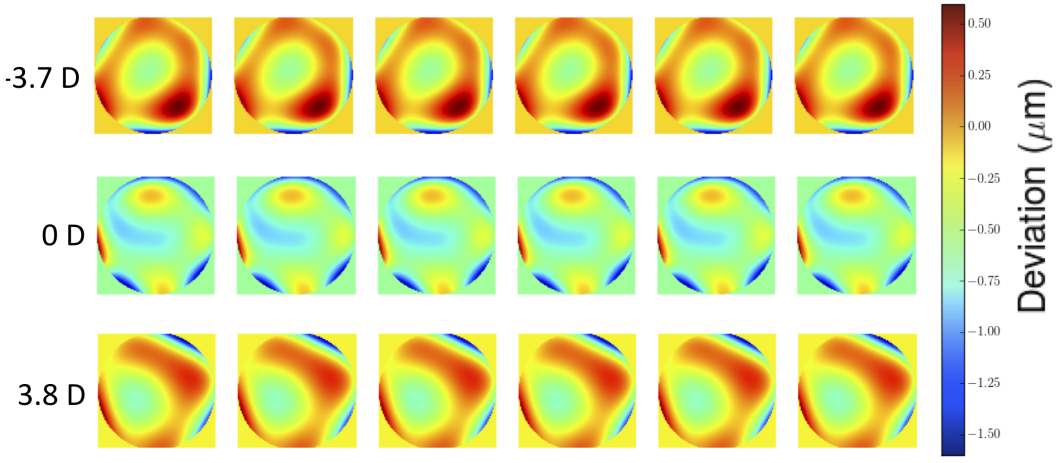

Figure. 5.3: Six repeated wavefront measurements showing variation over different refractive errors with a model eye. Only high-orders are shown.

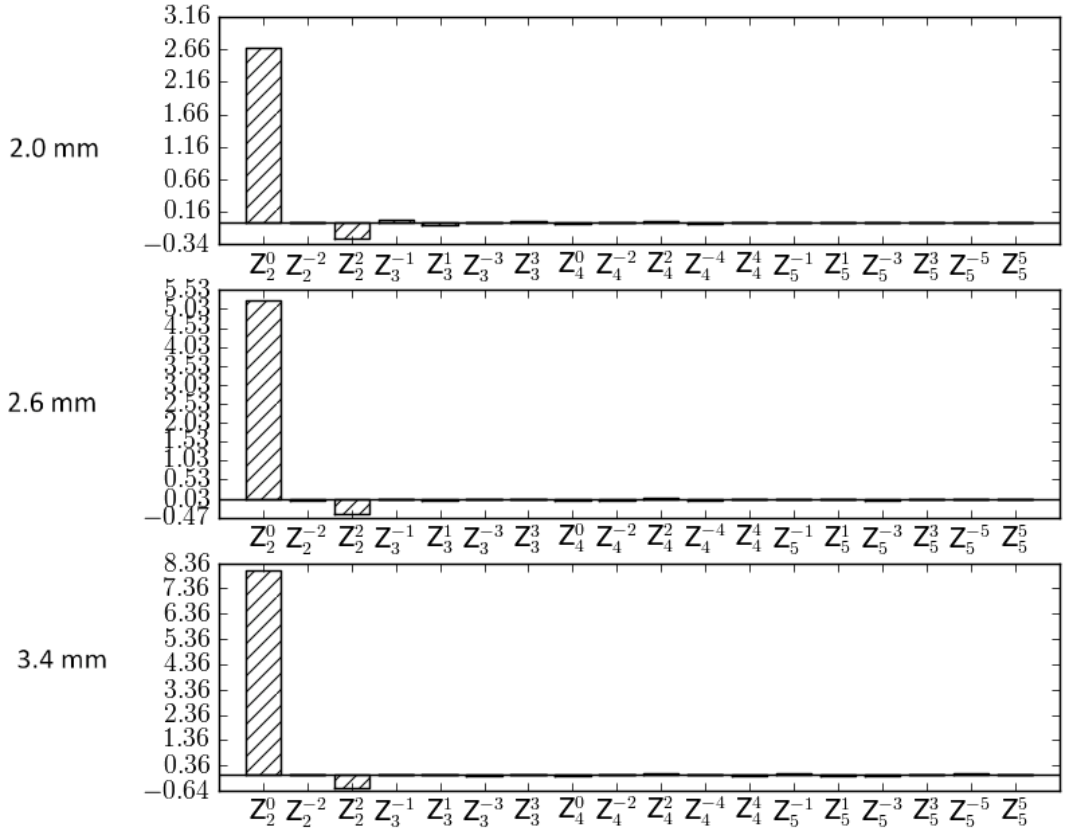

Figure. 5.4: A typical wavefront measurement for different pupil radii with a model eye.

5.4.1. Specifications of the model

The model eye consisted of stock lenses from Comar Optics mounted in a 25 mm lens tube. A schematic of the model is given in Figure 5.6. The first lens had a focal length that was half of the second lens, approximating the ratio of focal power of the cornea to the eye's lens. Between these two lenses an adjustable iris was inserted to simulate changes in pupil size. A space was left so that a phase plate could be inserted in the future but in these experiments a flat ($\frac{\lambda}{4}$) glass window was used. Beyond the second lens was an adjustable back which could be screwed in and out to simulate myopia or hyperopia. The model was constructed in two halves that were screwed together giving access to the glass window without disturbing the position of the other components. All components were held

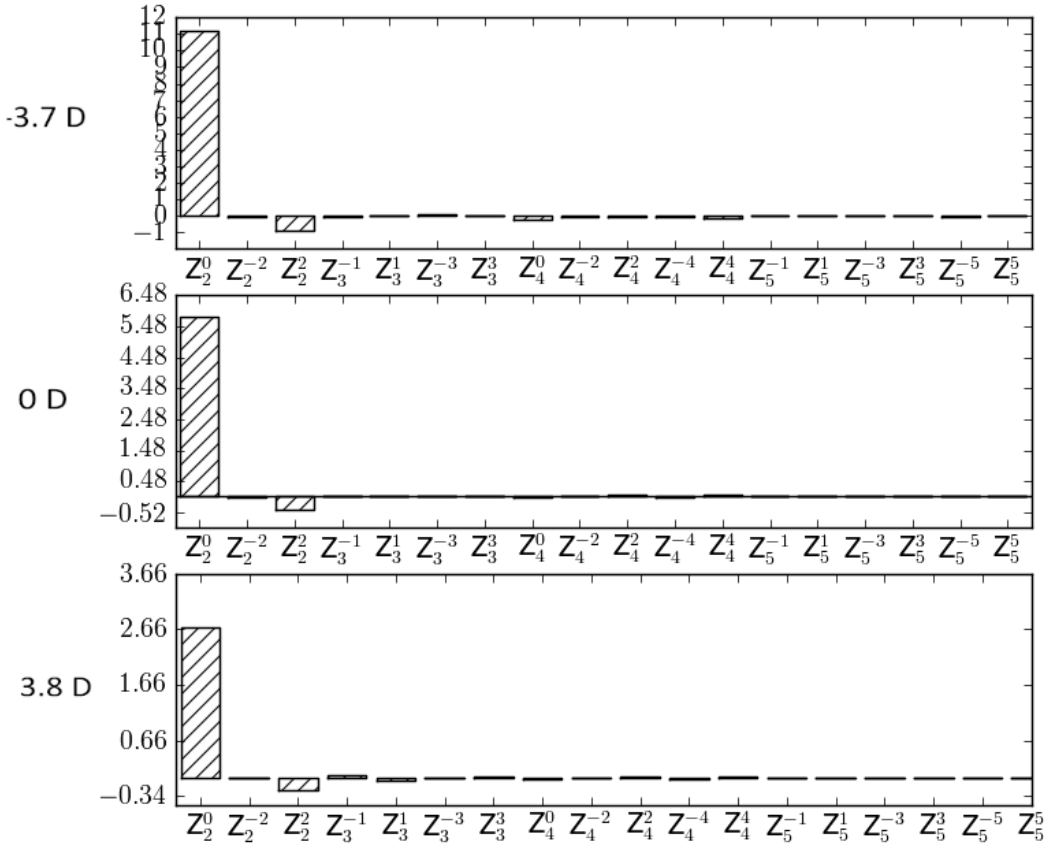

Figure. 5.5: A typical wavefront measurement for different refractive errors with a model eye.

with adjustable locking rings so that the model could be re-configured. The model was approximately a 2:1 scale model of a human eye as a 1:1 scale model would require lenses with such short focal lenses that there would not be enough room for an adjustable iris using stock parts. As a consequence the front surface of the model eye had a lower radius of curvature. Back reflections were therefore less divergent and so were visible when positioned in the Zywave. This was minimised by positioning the model eye at a slight angle to direct these reflections away from the wavefront sensor.

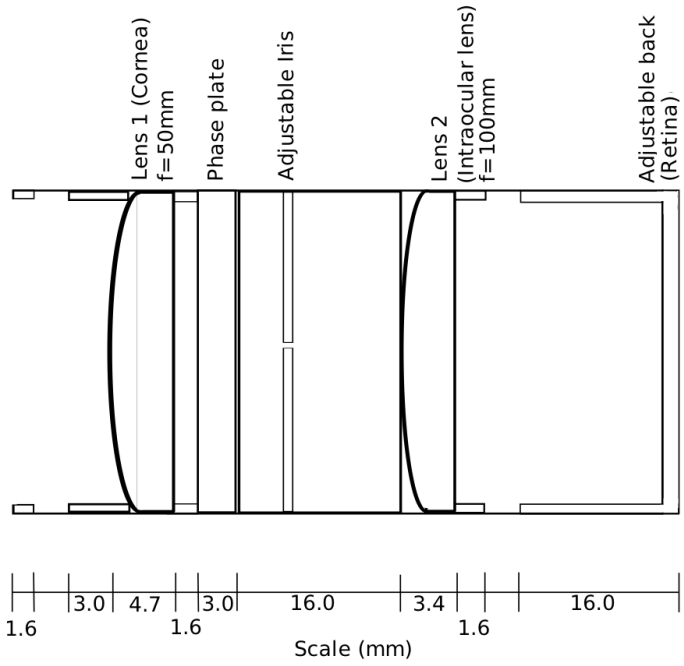

Figure. 5.6: Schematic diagram of the model eye used to check the repeatability of Zywave aberrometer measurements in the absence of subject-induced errors. This system is described in section 5.4.1. The scale at the bottom shows spacings in mm.

5.4.2. Variations with operator and realignment

For this experiment two operators took two sets of six repeated measurements of the aberrations of the model eye. One set was taken with realignment between measurements and the other was taken without. Operator 1 had more experience with the device and had trained operator 2 to use it. Operator 1 measured the pupil radius to be 3.05 mm (SE= 0.023 mm) in the case where the device was realigned between measurements and 3.15 mm (SE = 0.001 mm) when it was not. Operator 2 measured the pupil radius to be 3.06 mm (SE = 0.003 mm) in the case where the device was realigned between measurements and 3.07 mm (SE = 0.001 mm) when it was not. The standard deviations over 6 repeated measurements were compared with a 2 (operators) x 2 (alignment conditions) two-way repeated measures ANOVA (paired across Zernike modes). There was a significant effect of realigning the instrument on the standard deviation of the measurement ($F(1,51) = 17.256$, $MSE < 0.001$,

$p \leq 0.001$). The effect of realigning the instrument was significant for both operator 1 ($t(17) = 5.8$, $p \leq 0.028$) and operator 2 ($t(17) = 6.5$, $p \leq 0.021$). Another 2 x 2 two-way repeated measures ANOVA showed that there was no significant effect of realigning the instrument on the coefficient of variation ($F(1,51) = 0.267$, MSE = 0.163, $p = 0.612$). This suggested that the increase in standard deviation was accompanied by an increase in the mean. There was no significant effect of the changing the operator on the coefficient of variation ($F(1,51) = 0.340$, MSE = 0.194, $p = 0.567$) or on the standard deviation ($F(1,51) = 0.675$, MSE < 0.001, $p = 0.423$). Figure 5.7 summarises these results which show that realigning the instrument increased the standard deviation of repeated measurements by up to a factor of 4. This is in good agreement with the findings of Cervino *et al.* (2007) and is twice as large as that measured by Cheng *et al.* (2003). We conclude that the repeatability error of the Zywave is independent of the operator (assuming they have been trained) but is dependent on the ability of the operator to align the device, at least in terms of the standard deviation of repeated measurements. As there was no significant interaction between the operator and the alignment of the instrument ($F(1,51) = 0.1$, MSE < 0.001, $p = 0.773$ for the standard deviation and $F(1,51) = 0.3$, MSE = 0.445, $p = 0.362$ for the coefficient of variation) we suggest that the operators were equally capable of aligning the device when using a model eye.

5.4.3. Measurement of system-induced repeatability errors

By considering only data where the device had not been realigned we can quantify the system-induced measurement variance. Figures 5.8 and 5.9 show the standard deviations and coefficients of variance for each Zernike mode for both operators when the system was not realigned between measurements. These results show that there are small variations in the standard deviation between Zernike modes and large variations in the coefficient of variation. Again, differences in the coefficient of variation are due to differences in the mean amplitude. Our results show a lower standard deviation of repeated measurements

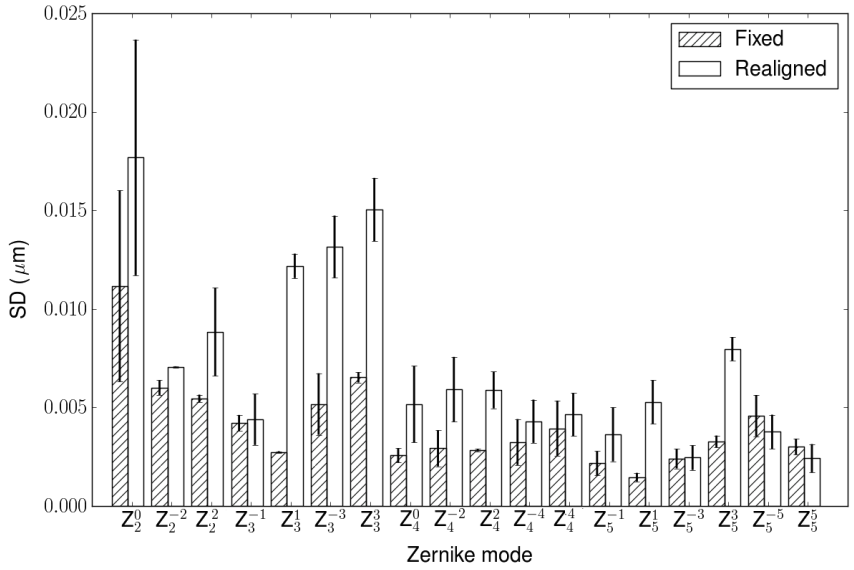

Figure. 5.7: Absolute repeatability error associated with realigning the Zywave, averaged over two operators (there was no significant effect of the operator on the standard deviation). Error bars represent the standard error on the mean. Realigning the instrument between measurements increased the standard deviation by up to a factor of 4.

(by a factor of approximately 2) than those found by Mirshahi *et al.* (2003), although the coefficient of variation is comparable. This suggests that the amplitude of aberrations measured by Mirshahi *et al.* (2003) were higher and this may have had an effect on the absolute error. Our results show that although the standard deviation was similar for higher-order aberrations and astigmatism, the coefficient of variation differed between types of aberration. Coefficients of variation below 5% were found for astigmatism, coma and spherical aberration however most other types of aberration had a relative error greater than 10% and some were greater than 100%, which is clearly unacceptable.

5.4.4. Variations with changing pupil size

The model eye was measured (without realignment) with pupil radii between 2 and 4 mm. The standard deviation and coefficient of variation were calculated for each Zernike mode, shown in Figures 5.10 and 5.11. The standard deviation was higher for low order

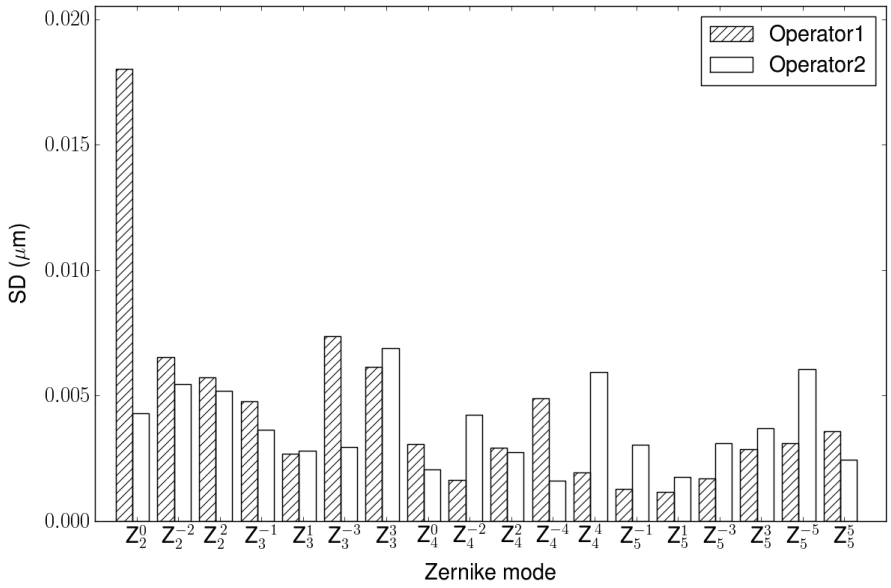

Figure. 5.8: Standard deviation of Zernike coefficients measured with the Zywave by two operators.

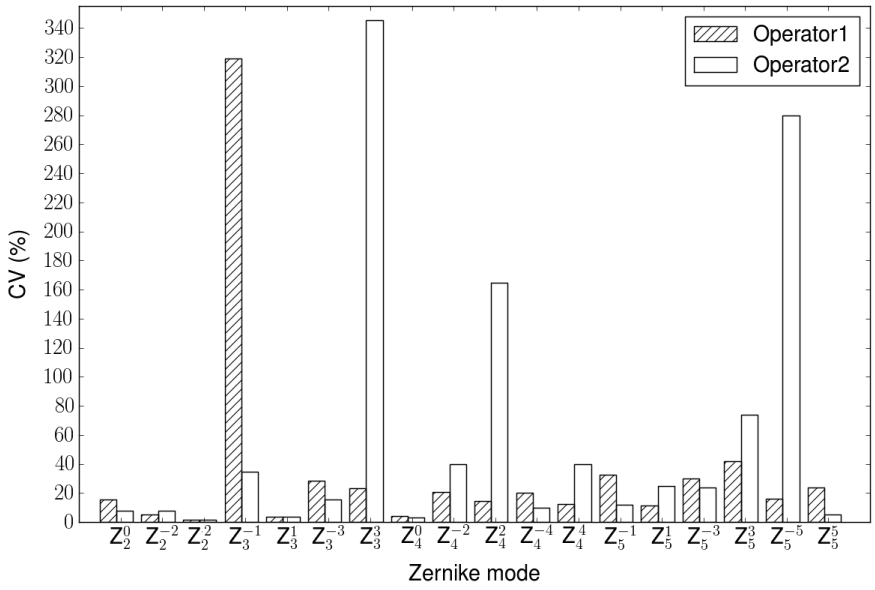

Figure. 5.9: Coefficient of variation of Zernike coefficients measured with the Zywave by two operators.

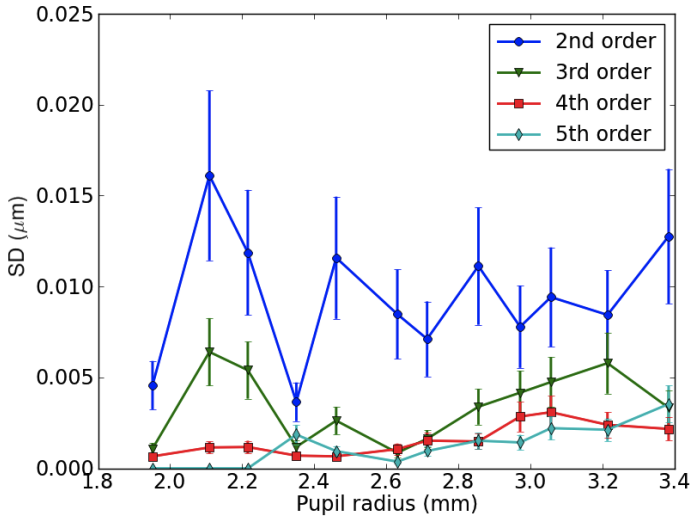

Figure. 5.10: Standard deviation of Zywave measurements with changing pupil radius, showing 2^{nd}, 3^{rd}, 4^{th} and 5^{th} order aberrations separately. For small pupil radii the amplitude of 5^{th} order aberration was zero since they are under sampled by the wavefront sensor and so the standard deviation was also zero. Error bars represent the standard error.

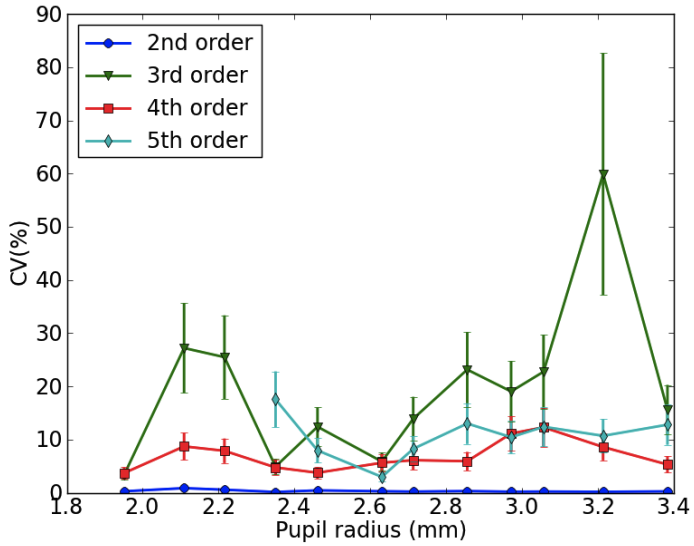

Figure. 5.11: Same format as Figure 5.10 showing the coefficient of Variation of Zywave measurements with changing pupil radius. For small pupil radii the amplitude of 5^{th} order aberration was zero since they are under sampled by the wavefront sensor and so the coefficient of variation could not be calculated. Error bars represent the standard error.

aberrations whereas the coefficient of variation was lower for second order aberrations and highest for third order aberrations. A 9 (radii) x 4 (orders of aberration) two-way repeated measures ANOVA showed there was a significant effect of the pupil radius ($F(8,175) = 4.6$, MSE < 0.001, $p \leq 0.005$) and an effect of the order of the aberration ($F(3,175) = 125.1$, MSE $= 0.001$, $p < 0.001$) on the standard deviation of repeated measurements. The interaction was not significant, which indicated that the effect of changing the pupil radius was the same for each order of aberration. There were no significant effects on the coefficient of variation, although this is not unexpected since the amplitude of aberrations also changes with the pupil radius. The change in standard deviation with pupil size could be affected by many factors. The pupil radius determines the number of Shack-Hartmann spots over which centroid estimation errors are averaged. The centroid error from any partially illuminated subapertures will be higher due to PSF elongation and a lower signal to noise ratio. Therefore the error on the spot position may not be uniform across the pupil and this may also change as the pupil size changes. However, without access to the optical components of the device and the software that calculates the wavefront measurement it is impossible to analyse these effects and how they propagate through the reconstruction process.

The Zernike coefficients that are measured vary as the pupil radius changes but we can account for this. As described in Chapter 2, a wavefront is defined as:

$$W(r,\, \theta) = \sum_{n=0}^{n_{\max}} \sum_{m=-n}^{n} c_n^m Z_n^m(r,\, \theta). \tag{5.1}$$

We can express the 2 dimensional matrices as vectors giving the wavefront vector, $\mathbf{W}$, and the Zernike reconstruction matrix, $\mathbf{Z}$. The coefficients (c) can be expressed as a vector of length n,

$$\mathbf{c} = \mathbf{Z}^{+}\mathbf{W}, \tag{5.2}$$

where $\mathbf{Z}^{+}$ is the pseudo-inverse of the Zernike reconstruction matrix. To account for differ-

ences in pupil radius the Zernike coefficients can be calculated using a central smaller radius of the Zernike reconstruction matrix and the wavefront vector. Figure 5.12 shows that this reduces the absolute error of repeated measurements.

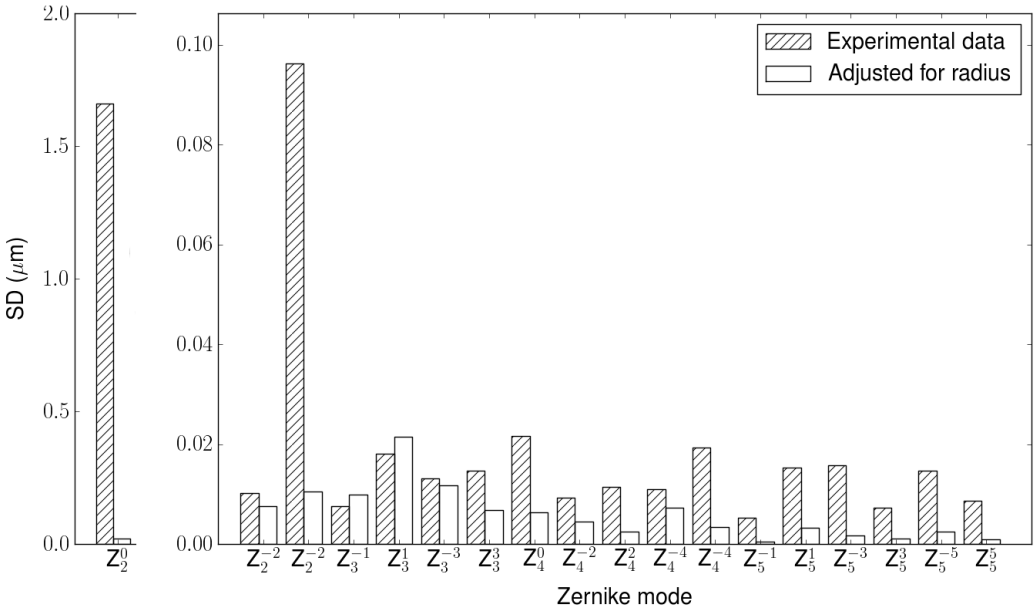

Figure. 5.12: Standard deviation of Zernike coefficients over pupil radii given in Figures 5.10 and 5.11 with and without compensating for pupil size. Wavefronts measured with larger pupil radii were cut to the smallest radius and Zernike coefficients were recalculated. Standard deviations improved for all Zernike modes except Z_3^{-1}, Z_3^1 and Z_3^{-3}. This shows that repeatability errors can be reduced by compensating for differences in pupil radius. The standard deviation of Z_2^0 was much larger and so is shown on a different scale.

5.4.5. Variations with changing focus error

The model eye was measured (without realignment) with focus errors between -3.8 and 3.7 D, achieved by moving the artificial retina backwards and forwards. This was done in order to simulate the effects of myopic or hyperopia on the system. It also gives an indication of the repeatability errors associated with changes in accommodative state, although this is not completely representative of a human eye's accommodative response since the optical

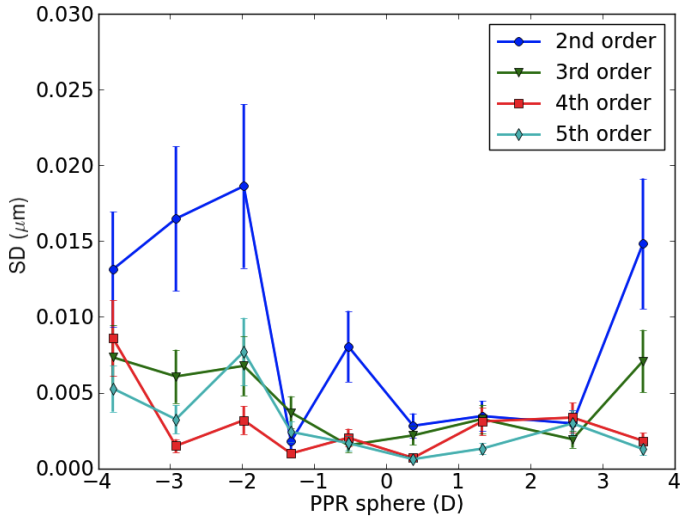

Figure. 5.13: Standard deviation of Zywave measurements with changing focus error, showing 2^{nd}, 3^{rd}, 4^{th} and 5^{th} order aberrations separately. Focus error error is expressed as the predicted phoropter refraction (PPR sphere) which is calculated by the Zywave. Error bars represent the standard error.

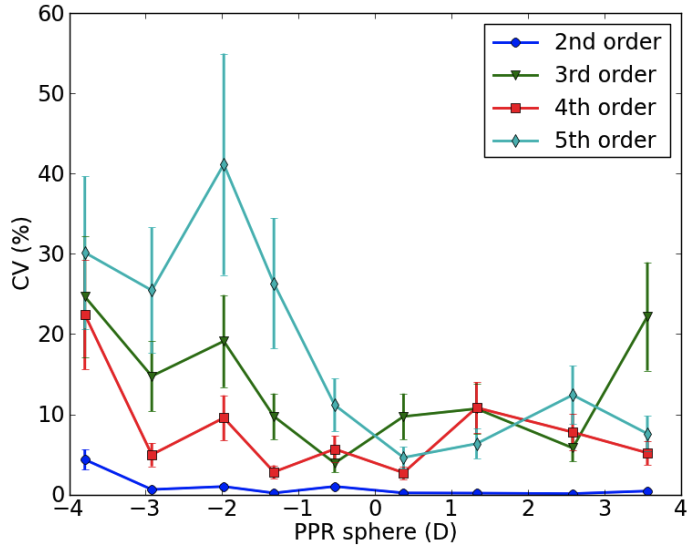

Figure. 5.14: Coefficient of Variation of Zywave measurements with changing focus error (as for Figure 5.13) showing 2^{nd}, 3^{rd}, 4^{th} and 5^{th} order aberrations separately. Error bars represent the standard error.

elements do not change. The standard deviation and coefficient of variation are shown in Figures 5.13 and 5.14. A 9 (focus conditions) x 4 (orders of aberration) two-way repeated measures ANOVA showed there was a significant effect of the focus error ($F(1.4,42.0) = 8.4$, MSE < 0.001, $p < 0.001$) and an effect of the order of the aberration ($F(3.0,42.0) = 20.4$, MSE $= 0.001$, $p \leq 0.002$) on the standard deviation of repeated measurements. The interaction was not significant, which indicated that the effect of changing the focus error was the same for each order of aberration. These effects on the coefficient of variation did not reach significance. As with changing pupil radius the absolute error is dependent on the focus error, but the relative error is not. We might expect this for defocus and spherical aberration since the amplitudes of these aberrations increase with increasing focus error. We do not expect the amplitude of higher-order aberrations to increase with focus error in a model eye since in the model eye the optical components are not changing. In this experiment the retina was translated by screwing the back of the model eye in and out and so there could be changes due to the difference in light back-scatter as the retina was rotated. The standard deviation and coefficient of variation were approximately symmetrical about 0 D. Variations with increasing amplitude of focus error could be due to the change in distance from the scattering surface to the pupil causing a change in the illumination profile. As described in Section 5.4.4 it is difficult to determine the effect this ultimately has on the wavefront measurement without access to the internal optical components.

5.5. Measurements on human subjects

Measurements were taken by Operators 1 and 2 of each other's eyes. Typical changes in the measure wavefronts and Zernike coefficients over 6 consecutive measurements are given in Figure 5.15. Changes over a day and within the same day are demonstrated in Figure 5.17. Since Section 5.4.2 has shown that both operators were equally repeatable in their alignment of the device we assume that differences between operators were not significant for this experiment. These were taken on 3 days and 6 repeated measurements were taken

in each of two sessions, one at 10am and the other 3pm. The order of eyes and operators was randomised to rule out effects due to fatigue. The measurements were compared with a 4 (eyes) x 2 (sessions) x 3 (days) x 18 (Zernike modes) repeated measures ANOVA. The ANOVA showed a significant main effect of the eye on the measurements ($F(1.1,183.2) = 301.5$, $MSE = 69.5$, $p < 0.001$) indicating that the four eyes were different over all Zernike modes. There was a significant effect of the Zernike mode on the measurement ($F(1.0,183.2) = 432.8$, $MSE = 1013.9$, $p < 0.001$) showing that the amplitudes varied among Zernike modes. There was also a significant interaction between the eye and the Zernike mode ($F(2.1,183.2) = 380.3$, $MSE = 1108.7$, $p < 0.001$) suggesting that each eye had a different pattern of amplitudes of the Zernike modes. Further results are discussed in the following sections.

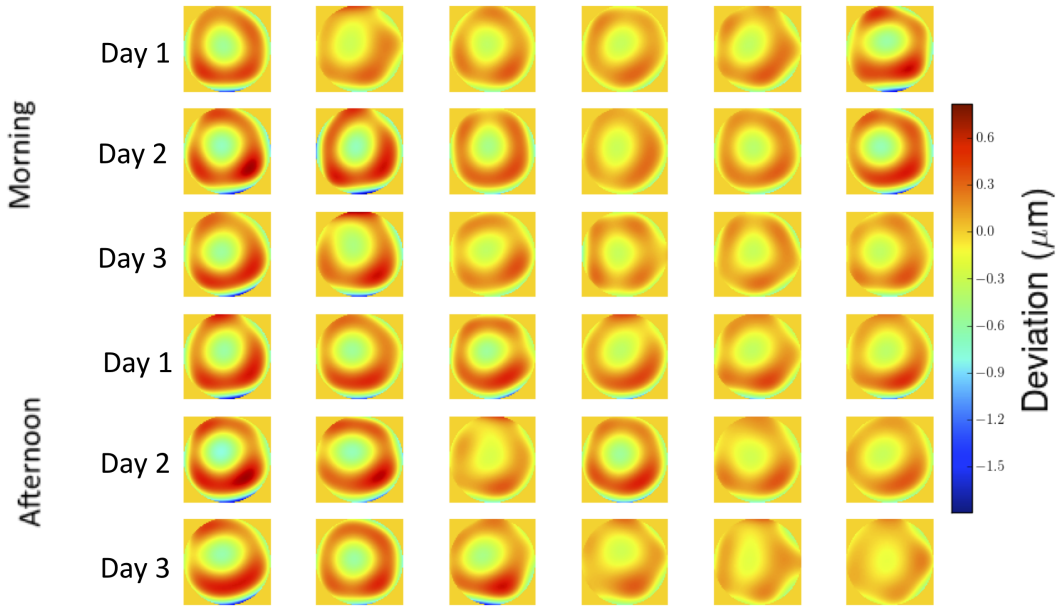

Figure. 5.15: Typical changes in the wavefront over 6 consecutive measurements. Only high-orders are shown.

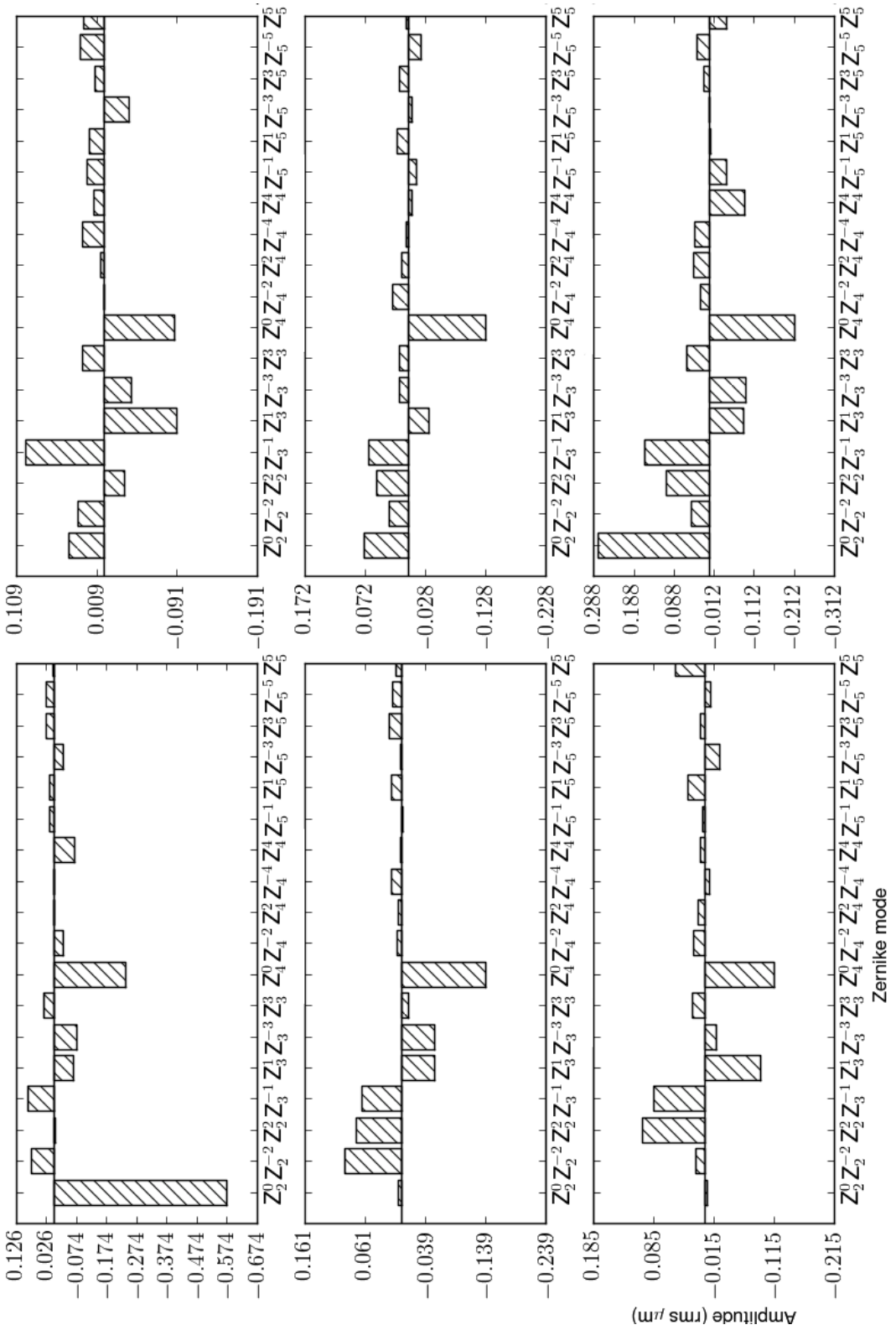

Figure. 5.16: Typical changes in the Zernike coefficients over 6 consecutive sessions.

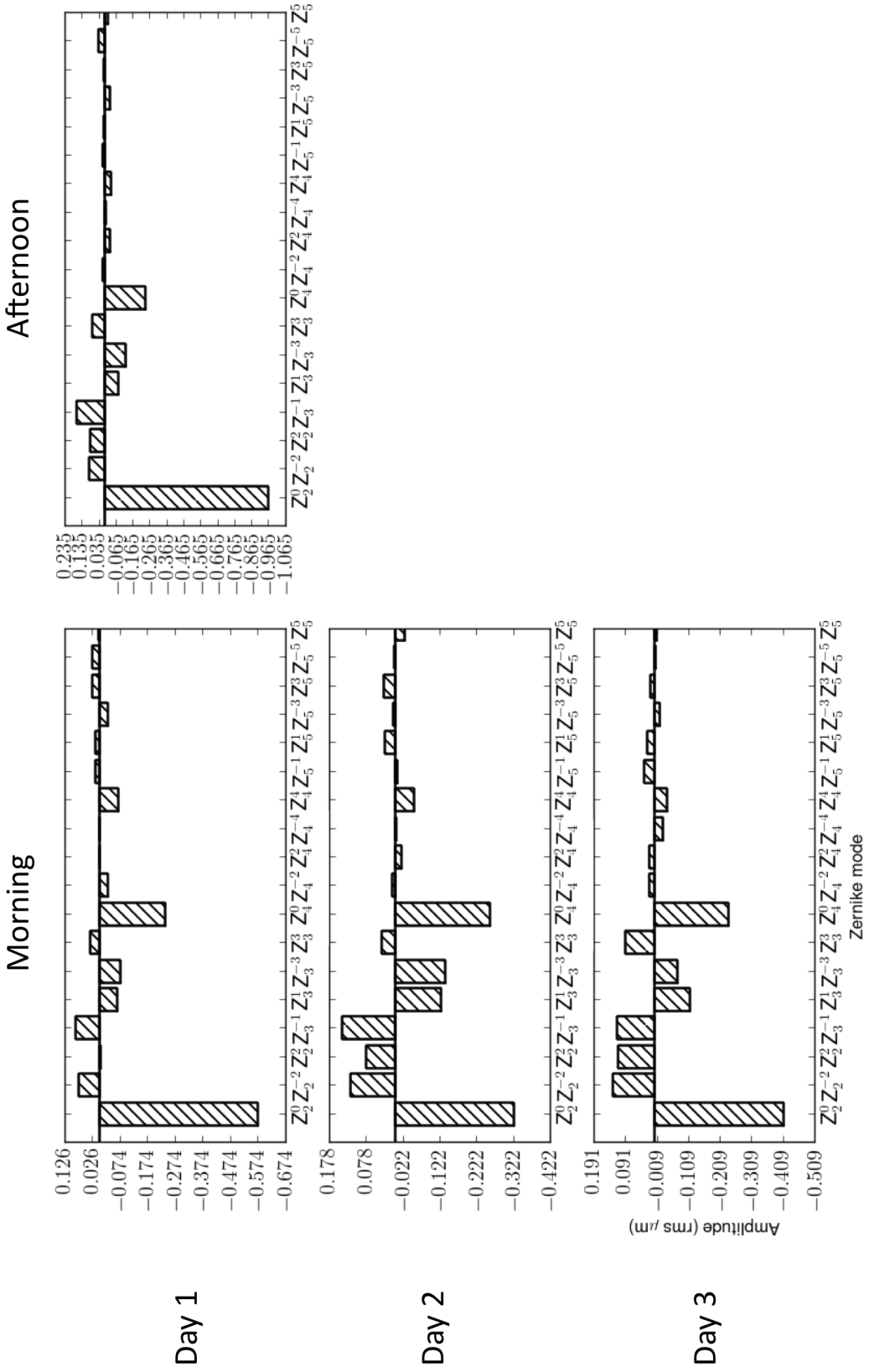

Figure. 5.17: Typical changes in the Zernike coefficients from day to day and between times of day.

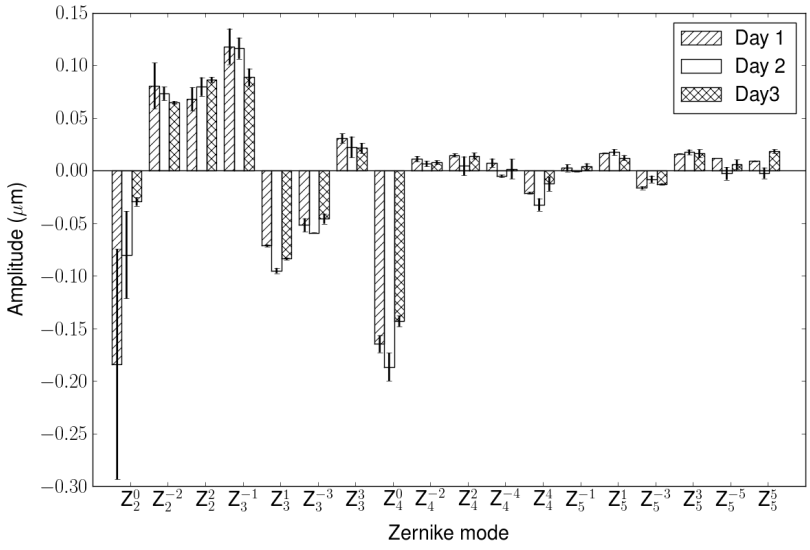

Figure. 5.18: Change in aberration measurement of LKY (right eye) over three days. Amplitudes are averages across sessions and the error bars represent the standard error on the mean.

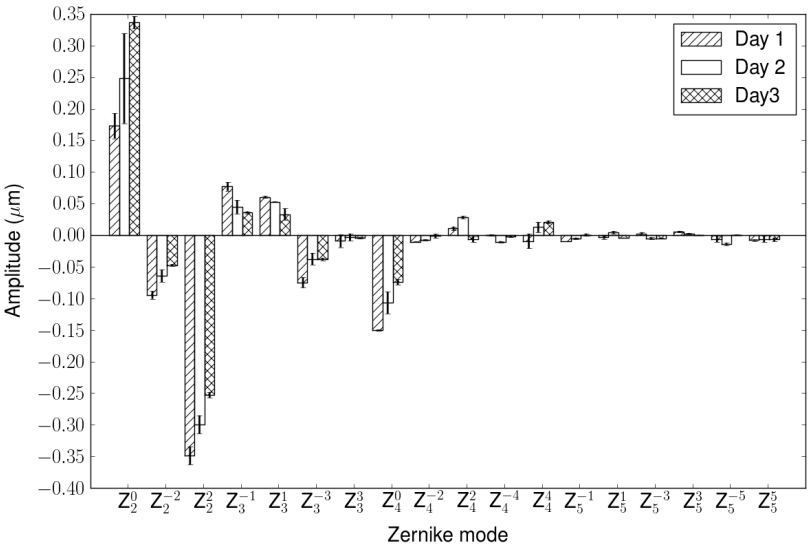

Figure. 5.19: Change in aberration measurement of LKY (left eye) over three days. Amplitudes are averages across sessions and the error bars represent the standard error on the mean.

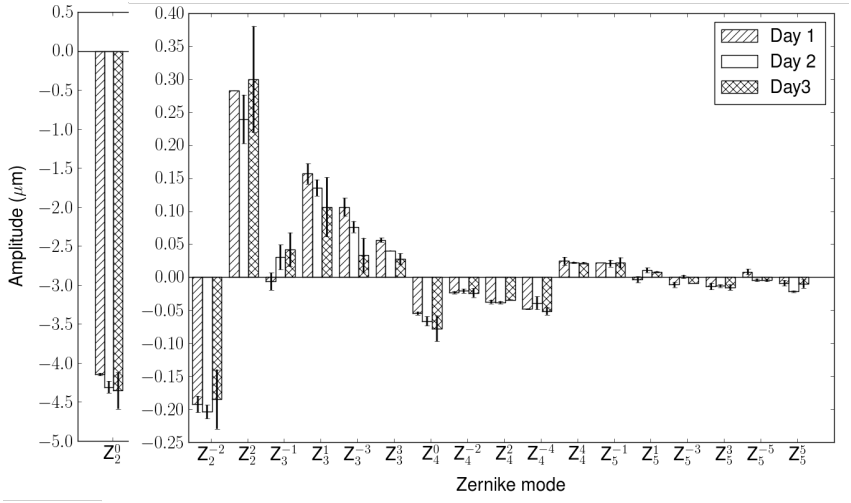

Figure. 5.20: Change in aberration measurement of SLP (right eye) over three days. Amplitudes are averages across sessions and the error bars represent the standard error on the mean.

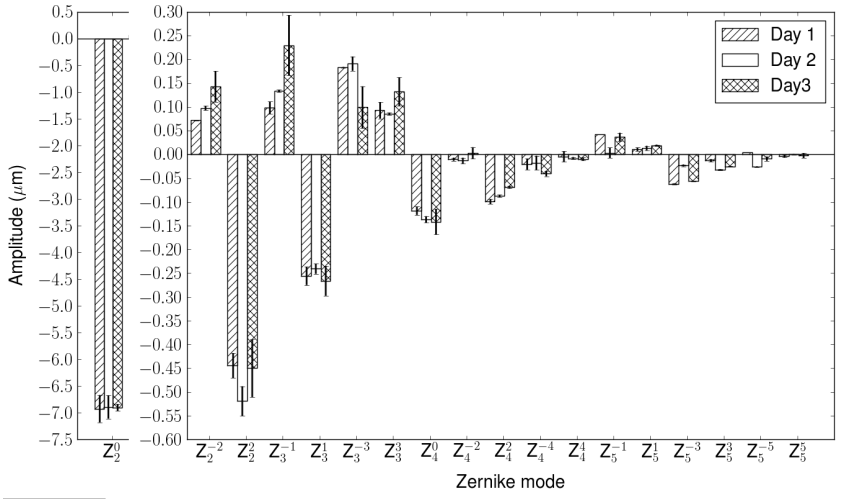

Figure. 5.21: Change in aberration measurement of SLP (left eye) over three days. Amplitudes are averages across sessions and the error bars represent the standard error on the mean.

5.5.1. Variations in the measurement from day to day

The main ANOVA showed no significant main effect of the day on the aberration measurement. This means over all Zernike modes and eyes the measurements made at the same time each day were not significantly different. These results are shown in Figures 5.18, 5.19, 5.20, and 5.21, where data for an individual day are averaged over sessions.

5.5.2. Variations in the measurement within the same day

Figures 5.22, 5.23, 5.24, and 5.25 show the aberration measurements at different times of day where data for an individual session are averaged over the three days. The main ANOVA showed a significant main effect of the session on the aberration measurements ($F(1.0,183.2) = 8.181$), MSE = 0.017, $p \leq 0.035$) and there was a significant three-way interaction between the eye, the session and the Zernike mode ($F(7.3,183.2) = 5.9$, MSE = 0.931, $p \leq 0.023$). This means that the Zernike coefficients varied not only between eyes and between sessions but that the difference between sessions varied with the eye being measured. A 4 (eyes) x 2 (sessions) x 18 (Zernike modes) three-way repeated measures ANOVA showed a significant effect of the session on the measurement for Zernike modes Z_2^2 ($F(1.0,65.1) = 37.6$, MSE = 0.049, $p \leq 0.002$) and Z_3^{-1} ($F(1.0,87.5) = 39.5$, MSE = 0.041, $p \leq 0.002$). There was a significant interaction between the eye and the session for Z_2^2 ($F(3.7,65.1) = 5.3$, MSE =0.026, $p \leq 0.028$). This means that the difference in the amount of (horizontal) astigmatism between sessions was different across eyes. A 4 (eyes) x 2 (sessions) two-way repeated measures ANOVA revealed a significant effect of the session for Z_2^2 in operator 2's left eye only ($F(1.0,25.0) = 62.9$, MSE = 0.111, $p \leq 0.001$).

From these results we can conclude that although there was a significant effect of the time of day on the aberration measurements this was driven by variations in the amount of astigmatism measured in one eye. Figures 5.22, 5.23, 5.24, and 5.25 show that there are differences between the two sessions. Although these differences didn't reach significance

for all Zernike modes it is possible that this was due to subject-, operator- and system-induced errors resulting in high variance in the measurements. Therefore the distributions of repeated measurements would not appear significantly different. We conclude that when making aberration measurements it is important to consider that the time of day may have a significant effect in some people.

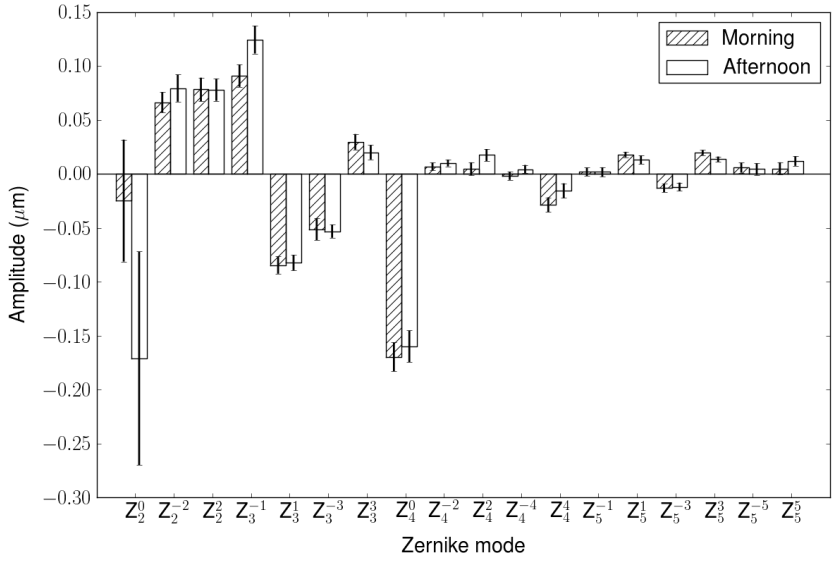

Figure. 5.22: Change in aberration measurement of LKY (right eye) within the same day. Amplitudes are averaged across days and the error bars represent the standard error on the mean.

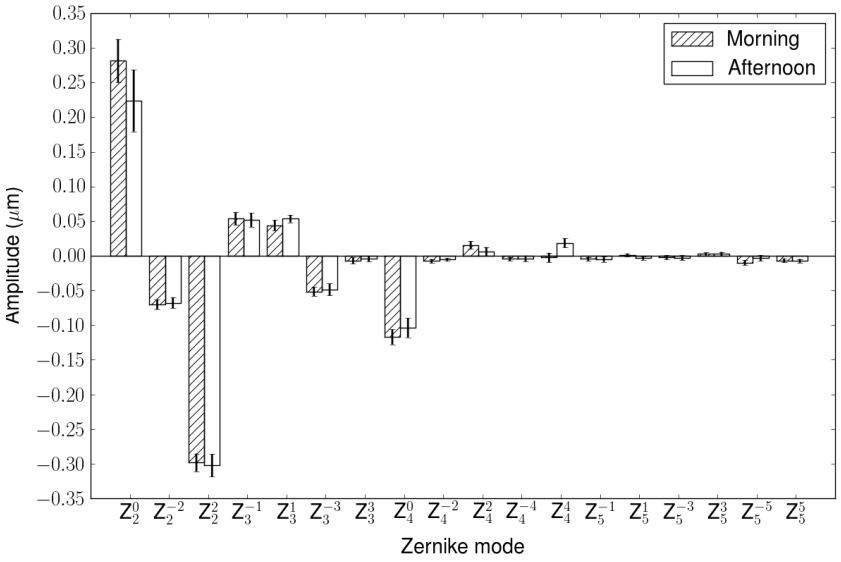

Figure. 5.23: Change in aberration measurement of LKY (left eye) within the same day. Amplitudes are averaged across days and the error bars represent the standard error on the mean.

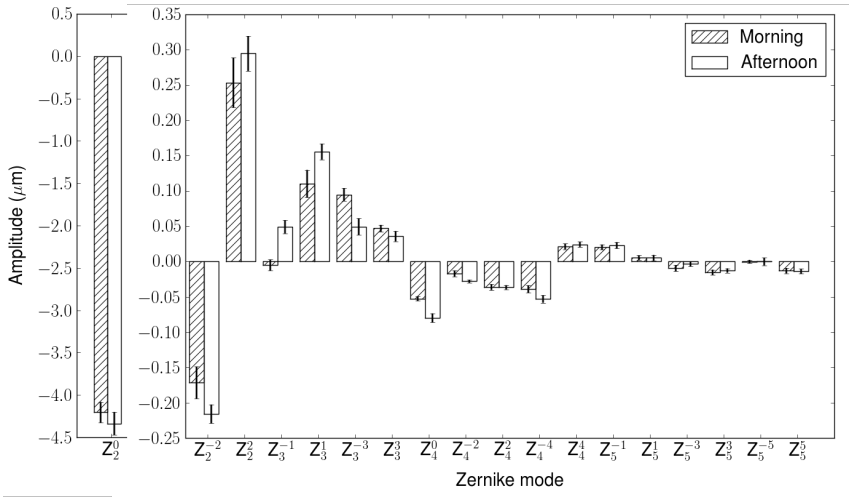

Figure. 5.24: Change in aberration measurement of SLP (right eye) within the same day. Amplituded are averages across days and the error bars represent the standard error on the mean.

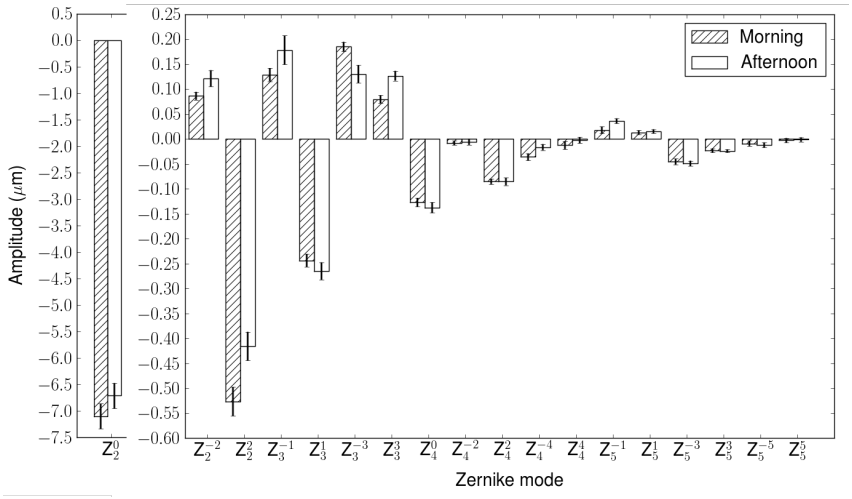

Figure. 5.25: Change in aberration measurement of SLP (left eye) within the same day. Amplitudes are averaged across days and the error bars represent the standard error on the mean.

5.6. Discusion

In this chapter we have shown that the Zywave aberrometer is not repeatable for all types of measurement. We find that system-induced repeatability errors, caused by errors in the centroid estimation for example, can be well above 100% for some aberrations. If these aberrations have a significant impact on visual performance then this error will be unsatisfactory when measuring eyes with the intent of producing a static correcting device. The difference between the actual aberration and the correction could be more than twice as large as the original measurement and this amplitude may be large enough to still impact vision.

We have shown that although the system-induced measurement error may be significant for some modes, there is a larger affect from changing the pupil size and the focus error. Third order aberrations are particularly affected by changing pupil radius, having relative repeatability errors of between 10 and 60%. Repeatability errors for fourth and fifth order aberrations are less affected by pupil radius, however all higher-order aberrations are af-

fected by focus error. Low order aberrations appear relatively robust to changes in pupil size and focus error. Errors in aberrations measurements can be reduced by recalculating Zernike coefficients over a fixed pupil radius.

We have demonstrated that the repeatability errors in aligning the instrument are operator independent (at least for the two operators we have tested). As long as the operator has been trained to use the device their ability to consistently realign the instrument does not improve with experience. There was a significant effect on the absolute error from this realignment suggesting that even though the repeatability is independent of which operator performs the measurement, it is still affected by human error.

Using measurements of the eyes of human subjects we have shown that aberrations of the eye tend not to change significantly from day to day within the limits of the instrument's repeatability. We know that the instrument repeatability is not always satisfactory, particularly when it must be realigned, and therefore there may be significant variations from day to day that we simply cannot measure. There was a significant effect of the time of day on the aberration measurement, although this was only significant for Z_2^2 measured in one eye. This tells us that we cannot always rely on a single set of measurements to give an accurate representation of the ocular aberrations in the long term. It is important therefore to make the ocular aberration measurement immediately before performing any psychophysical experiments on the subject. When producing a visual correction to be used in the long term many aberration measurements should be taken at different times and an average should be taken after accounting for pupil size and with outliers removed. However, this may still give an unsatisfactory result particularly if there is a large discrepancy in Zernike modes that impact vision significantly.

5.7. Conclusion

While the Zywave may produce repeatable measurements of low order aberrations this is not always the case for higher-order aberrations. The repeatability errors arise due to realignment of the instrument between measurements, changes in pupil size, changes in accommodative state (focus error) and system-induced errors such as in the centroid estimation. It is also possible that there are natural changes in the subject's aberrations over the course of the day. In the experiments described in this thesis we took three repeated measurements of the ocular aberrations of each eye of all of our subjects prior to testing. We believe that this improved the repeatability without increasing the experiment duration beyond a length that subjects could manage comfortably. We also accounted for the change in pupil size between the aberration measurement and the psychophysical experiment.

Chapter 6

Letter identification in the presence of an aberration

6.1. Introduction

There are several models of reading and word recognition, one of which is described in Chapter 3, and all of these include a letter recognition stage. Before discussing the effects of our chosen monochromatic aberrations on reading performance we will first describe their effects on letter identification. To study these effects we could use an acuity chart, as has previously been done (see Section 6.1.1 for details). However, these rely on changing the size of the letters, which in turn alters their spatial frequency content. As discussed in Chapter 3, letters are thought to be identified using a narrow band of spatial frequencies. The centre of this band is approximately 3 cycles per letter and has a bandwidth of approximately 1 octave (Solomon & Pelli, 1994). It is also thought that the centre of this frequency band can be shifted if the letters are blurred (Majaj *et al.*, 2002). Therefore, letter identification in the presence of different monochromatic aberrations, and in particular in comparison to the control condition, may be mediated by different spatial frequency channels. Furthermore, as the size of the letter is modulated these channels may or may not coincide with spatial

frequencies at which these aberrations cause phase changes. When they do coincide, spatial phase changes are likely to create spurious resolution which may introduce inconsistent features into the letters, causing them to be mis-identified. Under the argument made by Majaj *et al.* (2002), it is also possible that this spurious resolution drives the spatial frequency channel to a different part of the spectrum, causing these features to dominate over the correct features. To avoid introducing such ambiguity into our results we have chose to modulate the contrast of the letter, which is another commonly used clinical diagnostic for vision testing (Pelli, Robson, & wilkins, 1988), instead of its size.

In this chapter we discuss an experiment that investigated the contrast threshold for letter identification in the presence of an aberration. This aberration was either defocus, coma or secondary astigmatism, presented with a range of amplitudes of wavefront error. In Chapter 4 we implemented a cross-correlation analysis that predicted that defocus and secondary astigmatism would have a large impact on letter distinguishability, whereas coma would not. In this experiment we expect defocus and secondary astigmatism, but not coma, to cause a significant increase in the contrast required for letter identification.

In section 6.2 we explain the method by which the letter contrast was modulated and how the threshold value was determined. In section 6.3 we describe the stimuli we have used in this experiment and how they were presented to our subjects. The experimental method is detailed in Section 6.4 and the results are reported and discussed in Sections 6.5 and 6.7. In Section 6.6, we additionally relate our results to the cross-correlation analysis described in Chapter 4.

6.1.1. Previous studies on letter recognition in the presence of an aberration

Thorn and Schwartz (1990) tested the effects of dioptric blur on visual acuity and found the impact on Snellen acuity is greater than the impact on grating detection. They suggested that this was due to spurious resolution that caused the letters to become unrecognisable.

They found that there was a linear relationship between the amount of dioptric blur and the reduction in letter acuity. Akutsu *et al.* (2000) investigated the contrast and size thresholds for letter recognition in the presence of simulated dioptric blur. They attributed the effect of defocus on letter recognition to the limitation of useful spatial frequencies to those below the first minimum in the MTF (the first point at which the real part of the OTF is negative).

Oshika *et al.* (2006) measured the ocular aberrations of their subjects before testing their contrast and letter sensitivity functions and found that coma-like aberrations had a significant influence. Applegate *et al.* (2002) used rendered aberrations, as we have, to test the effects of higher-order aberrations on visual acuity and found that rms wavefront error is not a good predictor of visual acuity. They found the effect on visual acuity varied by up to 2 lines on a high contrast logarithm of the minimum angle of resolution (logMAR) acuity chart for different Zernike modes (Applegate *et al.*, 2002) and that those with low angular order (those near the centre of the Zernike pyramid) affect acuity the most. They also found a decline in visual acuity with increasing amplitude of aberration (Applegate, Ballentine, *et al.*, 2003). Li *et al.* (2009) investigated the effect of the amplitude of the aberration using an adaptive optics system in conjunction with a Freiburg acuity test. They also found that the reduction in visual performance was proportional to the rms wavefront error. Ravikumar *et al.* (2010) tested letter acuity in the presence of defocus, astigmatism, coma and spherical aberration (both individually and in combinations). They found that the 180° phase shifts in contrast caused by defocus produced the largest degradation in visual acuity, whereas the phase shifts introduced by coma ($< 180°$) had a smaller impact. They also note that the contrast in the phase reversed parts of the image is an important consideration as low contrast phase changes may not affect letter identification as much as high contrast changes.

Clearly there is a difference in the letter acuity (size threshold for letter identification) in the presence of both low and higher-order aberrations. As far as we are aware, there have been no studies that explicitly measure the contrast threshold for letter recognition as a function of the type and amplitude of higher-order aberrations, as we have in this

experiment.

6.2. Determining contrast threshold

A psychometric function describes some measure of performance in relation to some physical measure of a stimulus property. In this experiment we relate the probability of correctly identifying a letter to its contrast. When performance is expressed as a probability, the resulting psychometric function is typically S-shaped and is often described by a Weibull function (see Klein, 2001, for example). However, contrast is sensed logarithmically by the visual system so it is more appropriate to describe the psychometric function with a log-Weibull function, also known as a Gumbel function. The cumulative distribution function for the Gumbel distribution is:

$$P(X \leq x) = \gamma + (1 - \gamma - \lambda)(1 - exp[-10^{(x-\alpha)\beta}]). \tag{6.1}$$

where $P(X \leq x)$ is the probability that a real-valued random variable, X, will be found at at a value less than or equal to x. In the case of contrast threshold for letter identification measurements it is the probability that a letter will be correctly identified when its contrast is less than or equal to x. The parameter γ is the guess-rate (probability of a correct response due to chance), λ is the lapse-rate (probability of an incorrect response due to subject error), β is the slope of the function and α is threshold value for letter identification. Contrast threshold can be defined by a particular probability of a correct response or, as in this algorithm, it can be defined as the contrast value at which the slope is a maximum. The probability corresponding to the threshold value is that at $x = \alpha$, and therefore

$$P(X \leq x) = \gamma + (1 - \gamma - \lambda)(1 - exp[-1]). \tag{6.2}$$

In this experiment 26 different letters were presented and so a guess-rate of $\gamma = \frac{1}{26}$ and a lapse-rate of $\lambda = 0.01$ (Harvey, 1986) were used, giving a threshold response probability

of 0.64. To determine the threshold, the contrast is modulated with a staircase algorithm that calculates the contrast of the next trial based on the subject's responses to previous trials. This method assumes the probability of a correct response is a monotonic function of the stimulus level, that the responses are independent of each other and of the preceding stimuli and that the function doesn't change over the course of the experimental session. There are many types of staircase algorithm but the most efficient are those that use the maximum-likelihood technique (Pentland, 1980).

6.2.1. Maximum likelihood adaptive staircase procedure

In this experiment we used a staircase algorithm employing maximum-likelihood parameter estimation (ML-PEST) (Harvey, 1997). This algorithm was implemented using the Matlab Palamedes toolbox (Prins & Kingdom, 2009). This technique is based on the maximum likelihood techniques proposed by Hall (1968) applied to parameter estimation by sequential testing (PEST) (Taylor & Creelman, 1967; Findlay, 1978) and was developed by many others including Pentland (1980) and Watson and Pelli (1983). Using the ML-PEST technique, the experimenter predetermines a set of psychometric functions characterised by a range of values for the threshold (α) with fixed values for γ, λ and β. When a subject is presented with a stimulus of a given contrast, each of these candidate functions is used to predict the probability of a correct response at that contrast. This results in a likelihood function, describing the probability of a correct response as a function of the candidate threshold value. This likelihood function L_i is calculated at each new trial and is multiplied by the product of all of the previous likelihood functions, L, giving $L(L_i)$ for a correct response and $L(1 - L_i)$ for an incorrect response. The shape of the likelihood function evolves with each trial such that the centre of its peak shifts towards the most likely threshold value and the width of the peak narrows. The most likely threshold determined on a trial is used as the contrast in the next trial. Using this method the staircase algorithm attempts, at each trial, to present the stimulus with a contrast that is the best estimate of the thresh-

old value. Hence, this method is very efficient in terms of the number of trials required to determine an adequate estimate of threshold. Where other staircase procedures would attempt to calculate the threshold value by fitting a psychometric function to the responses over the whole session, ML techniques do not require the fitting of a psychometric function. ML-PEST assumes that the psychometric function has a specific mathematical form, but this is a reasonable assumption (Harvey, 1986).

6.2.2. Expected staircase performance

When initiating the staircase algorithm it is necessary to define a set of psychometric functions characterised by a threshold value. The values for γ, λ and β must fixed at some predetermined value. As described in Section 6.2, γ was set to $\frac{1}{26}$ and λ to 0.01. To choose the value of β, the slope, we examined data from preliminary experiments and found that a value of 30 (in units of increase in probability of a correct response per increment of contrast) for most of the aberrations we had tested. However, incorrect input parameters can cause the staircase to converge on the threshold value less efficiently. To ensure the algorithm converged sufficiently to give a reliable estimate of the threshold we decided to define a minimum acceptable value for the likelihood of that threshold being correct, based on the input parameters we have defined. The staircase's performance was investigated using different ideal observers (simulated observers that perform in an optimal way) that were blind to the identity of the letter and simply responded correctly or incorrectly with a predetermined probability distribution. This probability distribution was a Gumbel function in which the slope and threshold values could be varied to mimic different observers. We used ideal observers with different slopes and thresholds and recorded the corresponding error on the threshold value calculated by the staircase algorithm. For each ideal observer the staircase was repeated 1000 times and averages of the errors were calculated. Figure 6.1 shows the result of matching the input parameters to the ideal observer exactly and comparing the maximum likelihood value for a particular trial with the error on the threshold

estimate for that trial. This shows that for a maximum likelihood value above 0.4 we can expect an error of less than 5% on our threshold estimate, provided we choose the correct slope value. Figure 6.2 shows how this error changes as we change the slope value for our ideal observer. These data show that it is better to over-estimate the slope than under-estimate it. Although maximum likelihood values over 0.4 produce similar errors on the threshold estimate in simulation, in reality our subjects may make more mistakes than predicted. In the preliminary experiments we found that the maximum likelihood value typically reached at least 0.9. Therefore a value of 0.8 was expected to be conservative enough to allow for our subjects making unforeseen errors without unnecessarily rejecting too many threshold estimates. These errors may be due to typing errors or may relate to variability in the difficulty of the task, if for example some letters are more difficult to identify than others. Threshold estimates were calculated from the maximum likelihood value over 20 trials and this criterium was only used to determine if further trials were needed or if data should be rejected.

6.3. Stimuli

The stimuli used in this experiment were single lower case letters in courier font (the same as in the reading experiment). The contrast of these letters was modulated using a staircase that iteratively attempted to find the threshold for recognition. This algorithm is described in Section 6.2.1. We attempted to minimise the effects of the accommodative response, described in Section 4.4.2, by including a full-contrast box around the letter. Stimuli were presented on a CRT monitor with an average luminance 27 cd/m^2 (530 Td) and the output was gamma corrected. As with the reading experiment the stimuli were presented using the CRS Matlab toolboxes and the contrast was carefully modulated to maintain a resolution of 8-bits per gun, explained in Section 6.3.3.

Sixteen experimental conditions were investigated (three types of aberration, each with 5

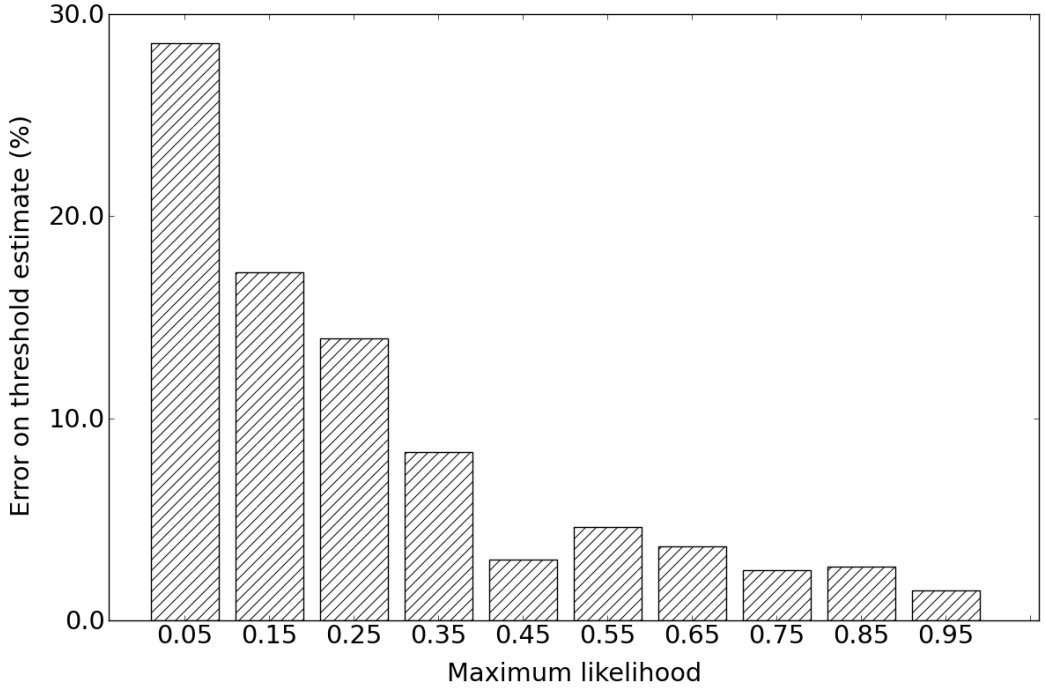

Figure. 6.1: Expected staircase performance as a function of the maximum likelihood value. An ideal observer, characterised by a psychometric function that matched the function used in the staircase algorithm, was tested 1000 times. The calculated threshold value was compared to the idea observer's pre-defined threshold value and the difference is expressed here as a percentage error. This was compared against the corresponding maximum likelihood determined in a given trial. These data show that the error on the threshold value was less than 5% for maximum likelihood values above 0.4.

different amplitudes of aberration, and a control condition). For each condition letters were presented at random and the probability of selection was weighted by the frequency counts of Jones and Mewhort (2004), as used in Section 4.3.1. The staircases for all experimental conditions were interleaved so that on any trial the type and amplitude of aberration was chosen at random.

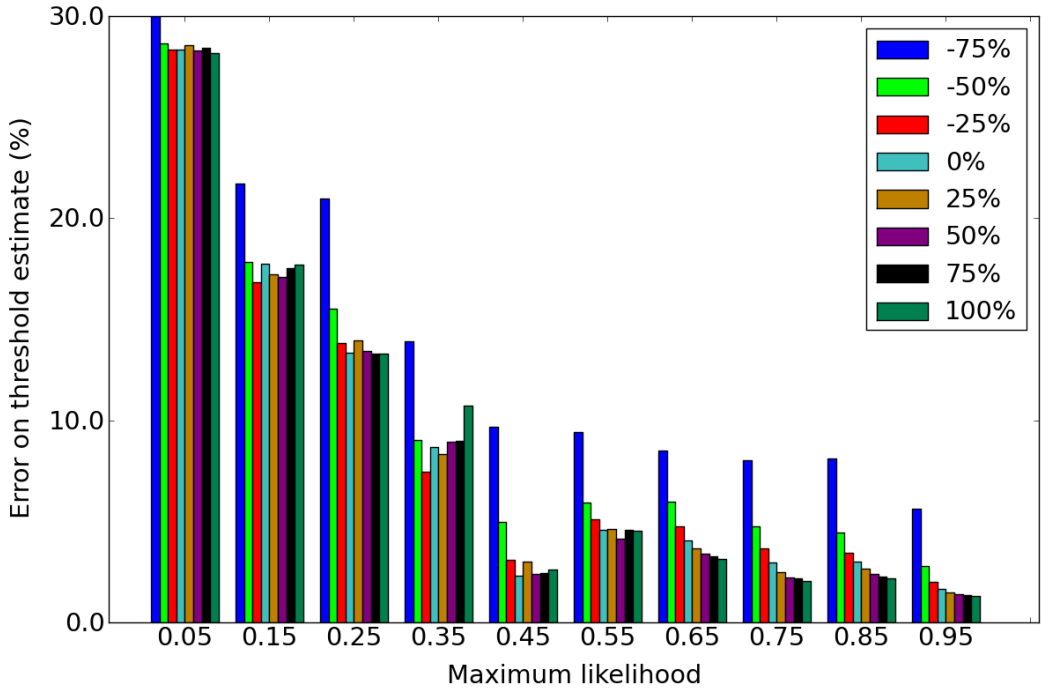

Figure. 6.2: Expected staircase performance as a function of the slope of the psychometric function that characterises an ideal observer, for a fixed slope used in the staircase algorithm to constrain the the psychometric function. The error on the slope is given as a percentage from the fixed value ($\beta = 30$). The results indicate the error on the threshold estimate for a given trial as a function of the corresponding maximum likelihood value. These data indicate that it is better to over-estimate the slope (positive percentage error) than under-estimate it. The error on the threshold is less than 10% over all slope values for maximum likelihood values over 0.4. However, if the error on the slope estimate is large (100% or more) it would be more appropriate to choose a maximum likelihood of 0.9 as our criteria for terminating the staircase.

6.3.1. Text size

We wanted the letters used in this experiment to be as large as possible so that our subjects' visual acuity did not limit their performance. Previous studies of the spatial frequency band used for letter recognition quote their results in either cycles per letter or cycles per degree. To simplify our analysis we have used letters that span 1°, which is equivalent to an acuity of 20/240 (worse than the limit for being considered legally blind). We also wanted our stimuli to be viewed at the greatest distance so that pixels could not be

resolved. The pixel size of the monitor used in this experiment was 0.37 mm (2.7 pixels/mm) and the pupil size during the experiment was 2.5 mm. Using the equation for the diffraction limited resolution of a circular aperture, θ,

$$\theta = 1.22\frac{\lambda}{D}, \tag{6.3}$$

the resolution of the human eye at the peak wavelength of the monitor (625 nm) over a 2.5 mm pupil, is 0.3 mrad. In order to just resolve pixels of size 0.37 mm the monitor must therefore be placed at least 0.89 m away from the subject. In this experiment subjects were positioned 1.7 m from the monitor.

6.3.2. Amplitude of aberration

To keep to a reasonable experiment duration, we chose to test five amplitudes of each aberration. We conducted preliminary experiments to find a suitable range of amplitudes to test. We found that the interval 0.5 - 0.9 μm (0.1 - 0.2 D of equivalent defocus) gave a significant variation in performance and in particular amplitudes above 0.9 were difficult to test reliably. These amplitudes were calculated over a 2.5 mm pupil and were appropriate for letters spanning 1°. While these values are not the same as those used in the reading experiment they do cover a comparable range in terms of image quality. Examples of the stimuli for each type and amplitude of aberration are given in Figure 6.3. We also note that in the main experiment the slope value, β, was set to 60, rather than 30, for 0.8 and 0.9 μm of defocus and secondary astigmatism to give the best chance of the staircase converging sufficiently.

6.3.3. Changing the stimulus contrast

Thresholds for letter recognition were found in this experiment by varying the contrast of the stimuli. The contrasts were calculated using an adaptive staircase procedure, outlined

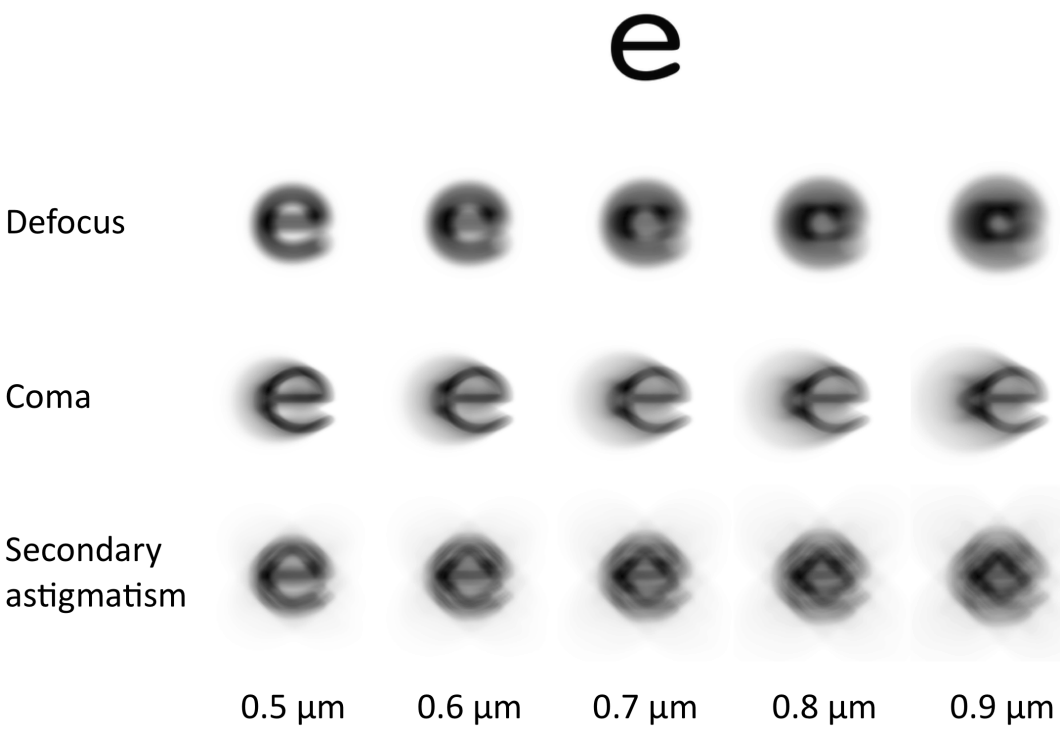

Figure. 6.3: Examples of the stimuli used in the letter experiment. Images are scaled such that they represent the size used in the experiment (1°) at arm's length (48cm).

in section 6.2.1. The CRS ViSage Stimulus generator was used to present the stimuli and control their contrast. The stimulus generator has a contrast resolution of 14-bits per gun but can only display 8-bits at once. These 8-bits are controlled with a colour palette which specifies what grey-level each value from 1 - 256 represents. To maintain a constant contrast resolution of 8-bits the colour palette was rescaled to give a letter with a chosen maximum and minimum grey-level. One value in the colour palette was reserved for black (1), which was used for the high-contrast border around the letter. The maximum grey-level was always white (256) and the minimum (other than black) was defined using the Weber contrast required on a particular trial. The Weber contrast is typically used for letters since it is appropriate for stimuli that consist of a small signal present on a large uniform background (the average luminance is approximately equal to the background luminance).

It is defined as:

$$\frac{l - l_b}{l_b}, \tag{6.4}$$

where l is the luminance of the signal and l_b is the luminance of the background. Assuming the monitor is gamma corrected there is a linear relationship between the gray-level and the luminance, l is therefore equivalent to the lowest gray-level in the letter and l_b is equivalent to 256. Once a suitable contrast for the next trial had been determined by the staircase the colour palette could be arranged to span l - 256 in 255 steps, leaving the first element in the palette as 1. Therefore the minimum contrast that could be correctly displayed by the stimulus generator was $\frac{1}{2^{14}} \times 255 = 0.015$.

6.4. Experimental method

6.4.1. Subjects

Five subjects participated in this study, three male and two female, with a mean age of 32 years (SD = 7 years). Three subjects did not require any vision correction for viewing letters on a monitor at a distance of 1.7 m, one subject wore contact lenses and one wore spectacles. All subjects were familiarised with the font prior to the experiment and were told that letters would appear with the frequency they normally occur in the English language. Wavefront measurements were taken of subjects' right eyes (the same eye with which the task was performed) and the average rms amplitude of higher-order aberrations was 0.006 $\pm$ 0.003 μm. Measurements of subject's higher-order aberrations are given in Table 6.1

6.4.2. Procedure

The stimuli in this experiment were presented monocularly using a patch over the left eye, at a presentation distance of 1.7 m. A fixed pupil radius was used to minimise the

Subject	Higher order (μm)	Higher order (excl. spherical) (μm)	Coma (μm)	Secondary astigmatism (μm)
GDL	0.065	0.059	0.071	0.039
HES	0.029	0.019	0.024	-0.011
LKY	0.103	0.080	-0.007	-0.009
RB	0.023	0.017	0.026	-0.006
RJL	0.067	0.068	-0.090	0.000(3)

Table 6.1: Table of measured higher-order aberrations for subjects participating in the letter recognition experiment. Values are calculated over the pupil diameter used in the experiment (2.5 mm). Results shown are the total rms amplitude of all higher-order aberrations, the total rms amplitude of all higher-order aberrations except spherical (which is affected by spectacle prescription), the amplitude of coma (Z_3^1) and the amplitude of secondary astigmatism (Z_4^2).

effects of the subject's aberrations on the retinal image, described in Section 4.4.1, and to increase the depth of field. This was achieved by positioning a 2.5 mm pinhole as close to the eye as possible and the subject was asked to align themselves in the headrest such that the letter was clearly visible. Since the pupil was not conjugated to the eye's pupil there was a reduction in the field of view, however this did not interfere with the ability of the subject to see the stimulus or keyboard.

A single letter was presented for 200 ms, which corresponds to typical fixation durations during reading. Each stimulus presentation was preceded by a high contrast fixation cross that directed the subject's gaze to correct location and prompted their attention and response. This was particularly useful when the contrast dropped below threshold as the subject may have not perceived the presence of a stimulus. Subjects were asked to report which letter they had seen using keyboard and audio feedback was given to indicate whether the response had been correct. All sixteen staircases were completed in a single session of 30 minutes duration so that comparisons between conditions were not affected by practice or familiarity. Subjects were given the opportunity to take regular breaks. The session was complete once each staircase had calculated the threshold with a maximum likelihood of at least 0.8, although in most cases this value was over 0.95. If this had not been achieved within 30 trials the best estimate was recorded but excluded from further analysis. Each

subject completed 8 sessions, providing 8 repeated measurements of performance for every combination of amplitude and type of aberration. The contrast threshold for letter recognition of a subject within a single session was determined for each experimental condition and the threshold for the control data was subtracted. This reduced inter-subject and inter-session variability. An average was calculated over all sessions for each subject before data were averaged over all subjects. Data were subjected to repeated measures ANOVAs. When sphericity could not be assumed (see Appendix 1), a Greenhouse-Geisser correction was applied for epsilon values below 0.75 otherwise a Huynh-Feldt correction was used. In the reports below we quote corrected degrees of freedom.

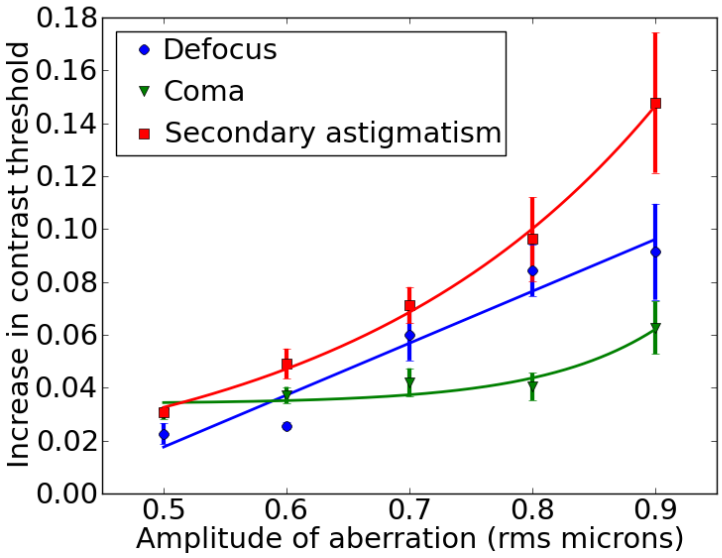

Figure. 6.4: Increase in the contrast threshold for letter identification from the control condition as a function of aberration amplitude. The mean contrast threshold in the control condition was 0.036 ± 0.002. Error bars represent the standard error on the mean, calculated over 5 subjects. Exponential growth curves have been fitted to the data via the least squares method. The growth constants were found to be 0.004 μm^{-1} (SSE < 0.001) for defocus, 10.6 μm^{-1} (SSE < 0.001) for coma and 3.9 μm^{-1} (SSE < 0.001) for secondary astigmatism.

6.5. Experimental results

The results of the experiment are given in Figure 6.4, which shows that secondary astigmatism had the largest impact on contrast threshold, followed by defocus and then coma. The growth constants suggest that defocus (0.004 μm^{-1}) caused an almost linear increase in the contrast threshold whereas secondary astigmatism (3.9 μm^{-1}), and particularly coma (10.6 μm^{-1}), did not. A 3 (types of aberration) x 5 (amplitudes of aberration) two-way repeated measures ANOVA revealed a significant main effect of the type ($F(1.1,41.3) = 23.716$, $MSE = 0.016$, $p \leq 0.007$) and amplitude ($F(4.0,41.3) = 19.695$, $MSE = 0.013$, $p < 0.001$) of the aberration. There was also a significant interaction between the type and amplitude of aberration ($F(8.0,41.3) = 3.244$, $MSE = 0.002$, $p \leq 0.008$) suggesting that the rate of increase in contrast threshold with the increase in aberration amplitude was different for the different types of aberration. We explored the significant interaction with one-way ANOVAs. There was a significant effect of the amplitude of aberration for defocus ($F(4,20) = 9.7$, $MSE = 0.005$, $p < 0.001$), for coma ($F(4,20) = 7.0$, $MSE = 0.001$, $p \leq 0.001$) and for secondary astigmatism ($F(4,20) = 9.4$, $MSE = 0.010$, $p < 0.001$). For each type of aberration the increase in its amplitude caused a significant increase in the contrast threshold. There was also a significant effect of the type of aberration at 0.5 μm ($F(2,12) = 6.0$, $MSE = 0.000$, $p \leq 0.025$), at 0.6 μm ($F(2,12) = 13.2$, $MSE = 0.001$, $p \leq 0.003$), at 0.7 μm ($F(2,12) = 5.6$, $MSE = 0.001$, $p \leq 0.030$), at 0.8 μm ($F(2,12) = 9.5$, $MSE = 0.004$, $p \leq 0.008$) and at 0.9 μm ($F(2,12) = 5.1$, $MSE = 0.009$, $p \leq 0.037$). At each amplitude of aberration the contrast threshold was different among the different types of aberration.

6.6. Comparison with confusion analysis

We compared our experimental data with the confusion analysis described in Chapter 4. Letters were compared via a cross-correlation with the same amplitudes and text size used in this experiment. The results are displayed in Figures 6.5, 6.6, 6.7 and 6.8. These results show that for each type of aberration there is a relationship between our predicted

measure of the confusability of letters and the data obtained in this experiment when comparing aberrated letters with both aberrated and unaberrated letters. Furthermore, these relationships are similar among the types of aberration, as demonstrated by Figures 6.7 and 6.8, which show the line of fit through all of the data points. We will demonstrate in Chapter 7 that this is not the case when considering reading performance rather than single letter recognition.

In this experiment letter recognition was tested at contrasts that were clustered around

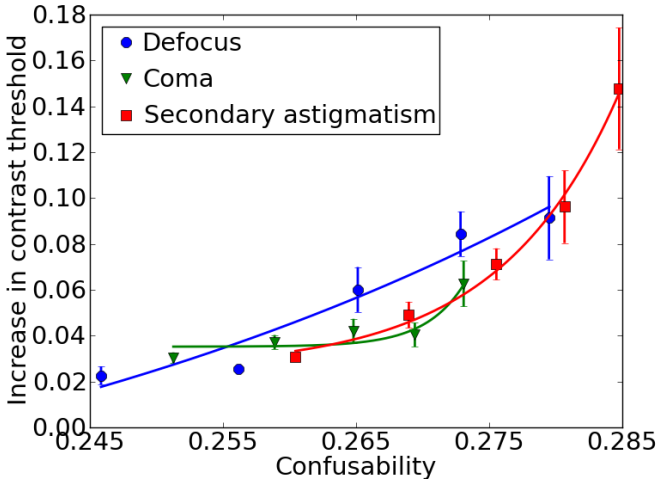

Figure. 6.5: Comparison between confusion analysis (comparing aberrated and aberrated letters) and the increase in contrast threshold for letter recognition, as shown in Figure 6.4. Data have been fitted with an exponential growth curves. The growth constant is 18 (SSE < 0.001) for defocus, 341 (SSE < 0.001) for coma and 112 (SSE < 0.001) for secondary astigmatism.

subjects' threshold. From these data we can also infer the particular letters that were easily confused, forming an experimentally determined confusion matrix. Figures 6.9, 6.10 and 6.11 show the probability of a subject incorrectly responding with a particular letter, with respect to the control condition (their normal guess rate). These results show that, for defocus, subjects tended to incorrectly respond with *c*, *l*, *n* and to a lesser extent *a* and *o* more often and with *r*, *s* and *t* less often. For coma subjects incorrectly responded more

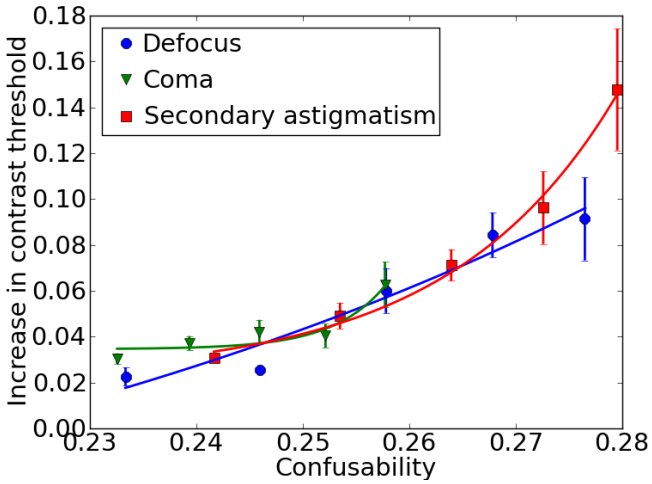

Figure. 6.6: Comparison between confusion analysis (comparing aberrated and unaberrated letters) and the increase in contrast threshold for letter recognition, as shown in Figure 6.4. Data have been fitted with an exponential growth curves. The growth constant is 12 (SSE < 0.001) for defocus, 203 (SSE < 0.001) for coma and 64 (SSE < 0.001) for secondary astigmatism.

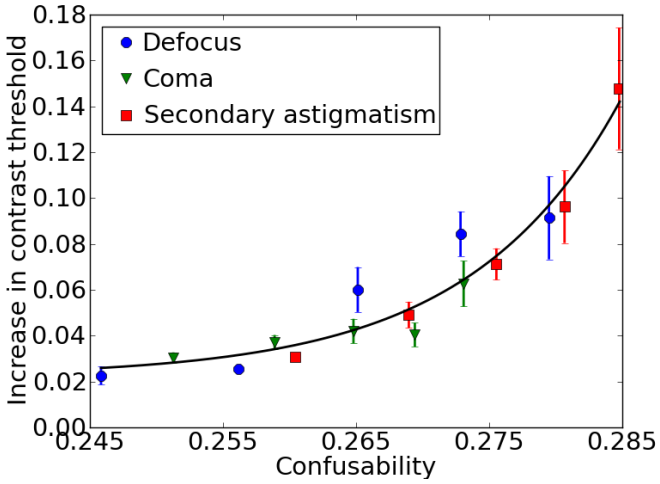

Figure. 6.7: As for Figure 6.5 with a single line fitted though all data points (growth constant is 89, SSE ≤ 0.001).

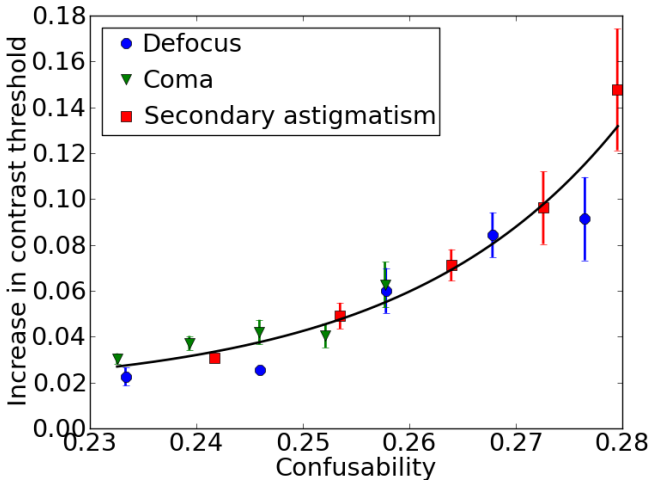

Figure. 6.8: As for Figure 6.6 with a single line fitted though all data points (growth constant is 50, SSE $\leq$ 0.001).

often with *c* and *r*. For secondary astigmatism subjects incorrectly responded with *a*, *c*, *n*, *o* and, at the highest amplitude, *x* more often and with *r*, *s* and *t* less often.

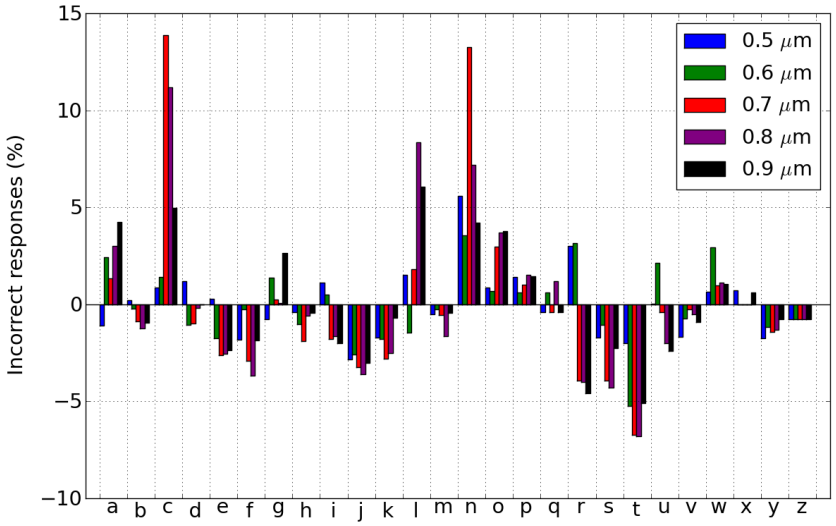

Figure. 6.9: Probability of subjects incorrectly responding with a particular letter, independent of what was presented, for defocus (probabilities sum to 1). Data are the difference from the control condition (the normal guess rate of of our subjects) for each of the five amplitudes of aberration tested.

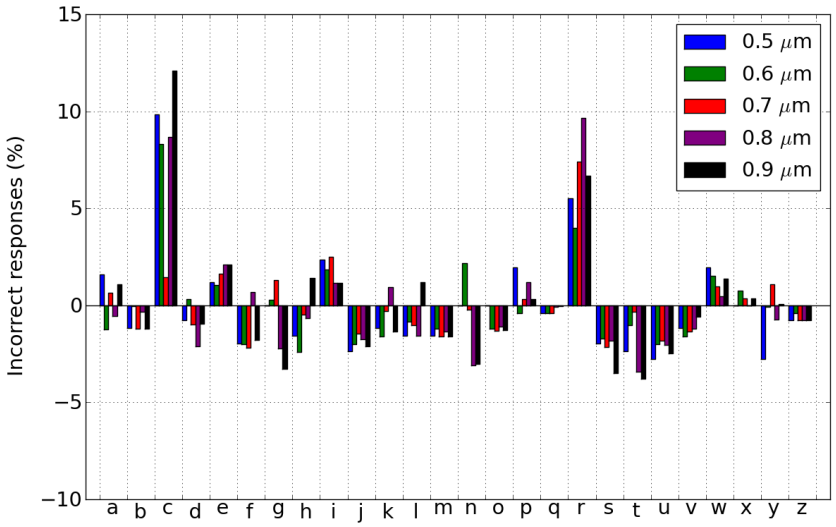

Figure. 6.10: As for Figure 6.9 but for coma.

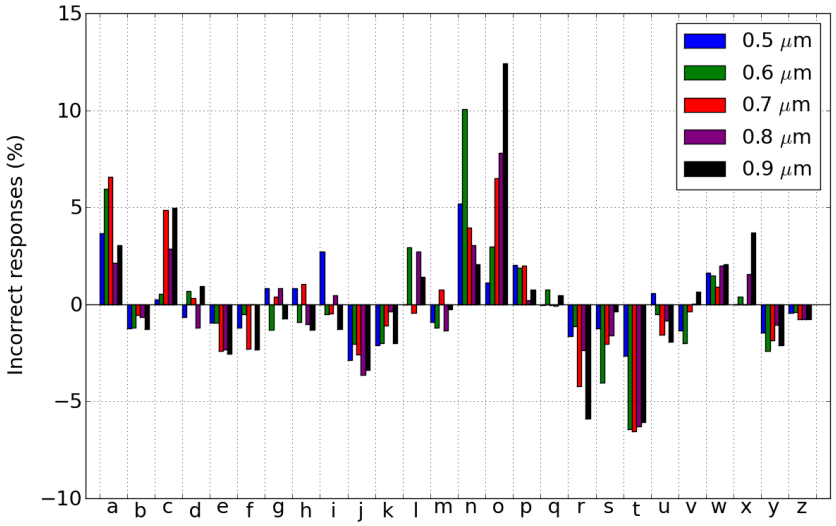

Figure. 6.11: As for Figure 6.9 but for secondary astigmatism.

6.7. Discussion

Recall that in Chapter 4 we predicted, based on cross-correlations between letters, that defocus and secondary astigmatism would have a large impact on letter distinguishability, whereas coma would not. The results obtained in this experiment support our hypothesis as defocus and secondary astigmatism impacted the contrast threshold for letter identification more than coma. The affect of coma on letter recognition was small up to an amplitude of 0.8 - 0.9 μm followed by a sudden increase that resulted in a large growth constant (10.6 μm^{-1}, SSE < 0.001). For defocus the contrast threshold increased by a factor of 4 from the smallest to the largest amplitude but the growth constant was small (0.004 μm^{-1}, SSE < 0.001), suggesting an almost linear trend in the data. Secondary astigmatism however caused a 7-fold increase in the contrast threshold and had a growth constant of 3.9 μm^{-1} (SSE < 0.001). Clearly secondary astigmatism had the most detrimental effect on letter identification over the range of amplitudes we have tested. We suggest that this could be due to the spatial phase changes caused by secondary astigmatism. The phase shifts caused

by defocus are either 0 or π, resulting in certain spatial frequencies being inverted. For secondary astigmatism however, the phase shifts are not just 0 or π and so spatial frequencies may be physically shifted with respect to one another. In either case these phase changes lead to spurious resolution and can create features which are inconsistent with the original form of the letter. This leads to the subject being unable to determine the letter and so guessing its identity, or worse, it results in them incorrectly identify one letter as another. This may be particularly problematic when the contrast is low if these incorrect features have a higher contrast than the correct ones. The features that should contribute to the correct form of the letter may be classified as noise when more prominent (but incorrect features) are present. An example of this is given in Figure 6.12. Defocus clearly causes spatial phase changes that alter the composition of the letters. In particular phase reversals caused by defocus lead the letter *m* to have two vertical bars and so appear almost like a letter *n*. Coma leaves the form of the letter relatively intact and causes a 'smearing' from right to left. Secondary astigmatism has a more drastic impact on the appearance of the letter causing a shadowing effect akin to double vision. In this case of the letter *v* it is not unreasonable to expect a subject to mistake it for a *w* or *x*.

We compared these data with our predicted model of performance, discussed in Chapter 4. We found that these was a correlation between our confusability value and the increase in contrast threshold for letter recognition. This relationship was similar for the different types of aberration and we found that a single relationship could be used to describe all three. We will show in the next chapter that this is not the case when measuring reading performance.

6.8. Conclusions

In this Chapter we have shown that the contrast threshold for letter identification increases with the amplitude of aberration and that the rate of increase is different for the

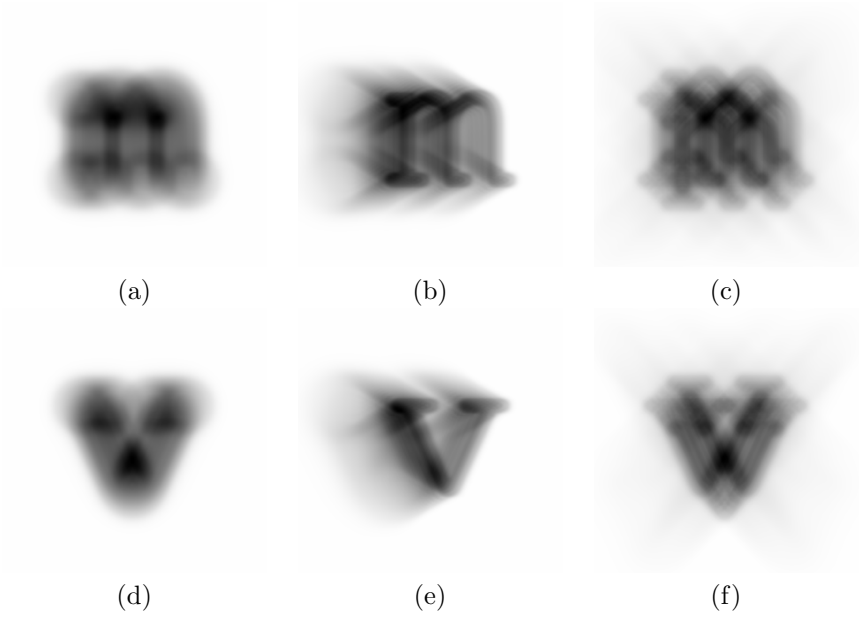

Figure. 6.12: Examples of the effect of spatial phase changes caused by 0.8 μm of (a) defocus, (b) coma and (c) secondary astigmatism on a the letter *m* and of 0.8 μm of (d) defocus, (e) coma and (f) secondary astigmatism on a the letter *v*. These letters are presented at a contrast value of 1. It is clear that both defocus and secondary astigmatism cause phase changes that significantly alter the form of the letter, but in different ways, whereas coma does not.

different types of aberration. Over the amplitude range we have investigated, secondary astigmatism impaired letter recognition the most, followed by defocus, with coma causing the smallest disruption to performance.

In Chapter 4 we explored the possibility of letters becoming less distinguishable in the presence of an aberration. Using the same analysis we compared our metric of confusion with the experimental results presented in this Chapter. We found that a relationship existed between our confusion metric and the increase in contrast threshold that could be described by an exponential function. In the next Chapter we will present the results of a study on reading performance in the presence of different aberrations.

Chapter 7

Reading performance in the presence of an aberration

7.1. Introduction

The main motivation behind this thesis was to understand the effects of different types of aberration on a higher-level visual task. We chose to study their effects on reading, which is an everyday visual task that we expect to be impaired by high amplitudes of ocular aberrations. In the previous chapter we examined the effect of defocus, coma and secondary astigmatism on letter identification. This was a prelude to studying reading performance, which involves letter recognition, but also incorporates effects on language processing. We hypothesise that the effect of aberrations on visual performance depends on the task. In this chapter we will show that this is the case even for related tasks, such as reading and letter recognition, due to effects on higher-level visual processing.

In this chapter we present eye movement data obtained during a reading task in which different rms amplitudes of defocus, coma and secondary astigmatism have been added to the stimuli. Performance measures over a full sentence, such as average fixation duration,

are reported and discussed in terms of overall impairment to reading. The impact of these aberrations on word identification specifically was investigated using fixation durations on predetermined target words. In this way, impairments to reading performance can not only be assessed in terms of the degradation of the visual input but importantly effects on linguistic processing can also be inferred.

Section 7.2 gives a detailed description of the stimuli used in this experiment and how they were constructed. The eye tracking system and the saccade detection algorithm used to analyse the data are outlined in Sections 7.4 and 7.3. The experimental method, results, discussion and conclusions can be found in Sections 7.6, 7.5, 7.7 and 7.8. The results in these sections will also be compared to the predicted performance obtained by our letter confusion analysis.

7.2. Stimuli

The stimuli chosen for this experiment were sentences limited to a single line when presented on the monitor. These sentences were constructed in a such a way that word-specific measures of performance (hence forward referred to as local measures) could be obtained in addition to overall performance (referred to as global measures). These local measures provide insight into effects on linguistic processes, as described in Chapter 3. This was accomplished by placing a target word in the middle of sentence and examining the difference in fixation times between types of target word.

Thirteen experimental conditions (three types of aberration, each at four different amplitudes, and a control condition) were investigated with four repeats of each, leading to a total of 52 trials. The construction of the sentences used in these trials is described in the following sections.

7.2.1. Sentence layout and presentation

A set of 52 sentence frames were constructed, in the middle of which a target word was embedded. This target word contained six letters and was chosen to have either a high or a low lexical frequency, determined by its Kučera-Francis written frequency (Francis & Kučera, 1985). As an example, the sentence frame "Sophie was trying to draw her xxxxxx but was finding it very difficult" had a target word ("xxxxx") that was either "sister" (high frequency, 38 occurrences per million words) or "ferret" (low frequency, 1 occurrence per million words). No target word was used in more than one sentence frame and subjects were presented with each sentence frame only once. This ensured that each complete sentence read by a subject was equally unfamiliar to them. For each experimental condition, four sentences were chosen such that two contained a high frequency word and two a low frequency word. These four sentences were chosen at random using a Latin square approach. This ensured that each sentence was presented with a different aberration type and magnitude from subject to subject, ruling out sentence-specific effects. The 52 sentences were presented randomly during the session to prevent subjects adjusting to a particular type and amplitude of aberration.

7.2.2. Sentence normalisation

The 104 sentences (2 lexical frequencies x 52 sentence frames) were pre-screened for plausibility, and the target words for predictability. This ensured that effects on reading performance could not be attributed to differences in sentence context. For example, a predictable target word might likely be skipped over or perhaps fixated for a shorter time compared to an unpredictable word, irrespective of the word's lexical frequency. Equally an implausible word might draw a longer fixation than a plausible one. This normalisation study was conducted using 65 subjects aged between 17 and 20, none of whom participated in the main experiment. For this test 112 sentences (2 lexical frequencies x 56 sentence frames) were constructed, allowing for some rejections. No sentence frame was observed

(a)

(b)

(c)

Figure. 7.1: An example sentence simulated with 0.3 μm of a) defocus, b) coma and c) secondary astigmatism. Only the beginning of the sentence is shown here. The full sentence read "Sophie was trying to draw her (sister/ferret) but was finding it very difficult" and was displayed on a single line. In this sentence the target word is either "sister" which is a high frequency word (38 occurrences per million words) or "ferret" which is a low lexical frequency word (1 occurrence per million words).

more than once by a single subject. Each sentence was assessed by 16 subjects on average.

Subjects were given the first half of 14 of the 56 sentence frames, up to but not including the target word. They were asked to predict the word that might likely follow next in the sentence, i.e. the target word. A sentence frame was rejected if either of its high or low frequency target words had been predicted by any single subject. This occurred in only two sentence frames.

Subjects were asked to assess the plausibility of 28 of the 112 full sentences by ranking each one on a scale of 1 (completely implausible) to 5 (completely plausible). After removing the two sentence frames rejected by the predictability test, the two surplus sentences with the lowest plausibility scores were rejected. These plausibility scores of the final 52 sentence frames are given in Figure 7.2. A paired t-test was performed on the average rankings of the final 52 sentence frames and confirmed that the high and low frequency target words were not significantly different in terms of their plausibility ($t(51) = 1.629$, $p = 0.113$).

The lexical frequencies of the target words in the final 104 sentences were significantly

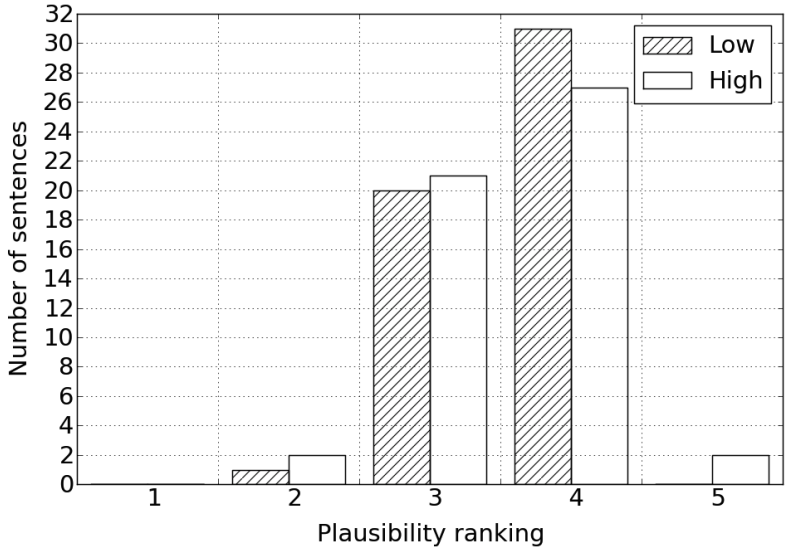

Figure. 7.2: Average plausibility rankings for the sentence frames with either the high or low frequency target word embedded. Results show no significant difference between high and low frequency words, confirmed by a paired t-test ($t(51) = 1.629$, $p = 0.113$).

different, as found by a paired t-test ($t(51) = 11.23$, $p < 0.001$). On average high frequency words had a frequency of 160 occurrences per million words (SD = 189 occurences per million words) and low frequency had on average 4 occurrences per million words (SD = 5 occurrences per million words). The lexical frequency distributions for the high and low frequency sets are given in Figure 7.3.

7.2.3. Text size

The 52 sentence frames had on average 71 characters that were required to fit on to a single line on a monitor. The angular size of these characters also needed to be well above the acuity limit of our subjects so that reading impairment was induced only by the aberration. The viewing distance was therefore limited by the requirement to maintain a minimum angular size of the characters and by the maximum character size achievable

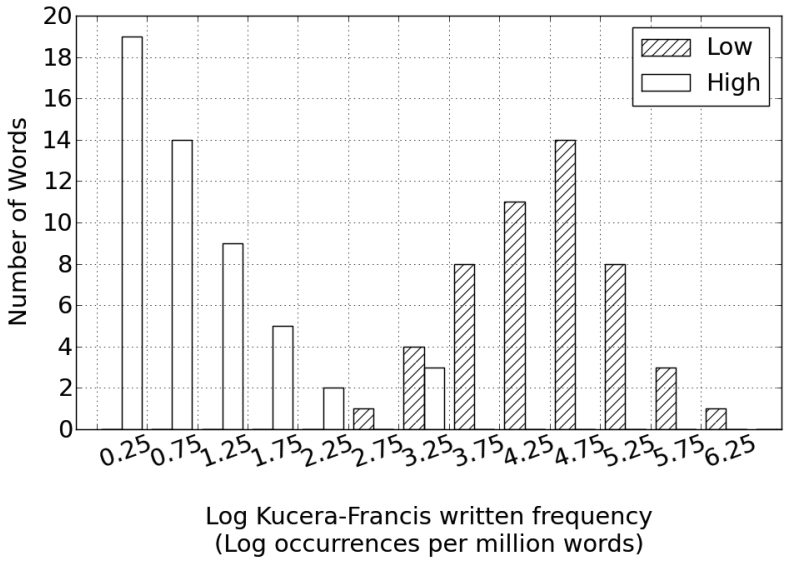

Figure. 7.3: Lexical frequency distributions for high and low frequency word sets, plotted on a log scale. These were found to be significantly different by a paired (unequal variances) t-test ($t(51) = 11.23$, $p < 0.001$).

whilst restricting all of the sentences to a single line. The distance between the subject and the monitor ideally needed to be large so that individual pixels could not be resolved. The pixel size of the monitor used in this experiment was 0.27 mm (3.7 pixels/mm) and the average pupil size during the experiment (tested prior to the experiment and later confirmed during the experiment) was 3.5 mm. Using Equation 6.3, which gives the diffraction limited resolution of a circular aperture, the resolution of the human eye at the peak wavelength of the monitor (625 nm) over a 3.5 mm pupil, is 0.2 mrad. In order to just resolve pixels of size 0.27 mm the monitor must therefore be placed 1.35 m away from the subject. However, the resolution of the human visual system is limited by the neural contrast sensitivity function as discussed in Chapter 3. The highest spatial frequency that can be detected by the visual system is 60 c/deg which is equivalent to 3437 cycles/radian. For a monitor with pixel scale, s, the presentation distance, d, which gives a spatial frequency, f, is given by:

$$d = \frac{f}{2s \tan \frac{\theta}{2}}. \tag{7.1}$$

The presentation distance for a monitor with a pixel scale of 3.7 pixels/mm is 0.85 m for a spatial frequency of 3437 cycles/radian. Using Equation 7.1 it is possible to re-scale the human contrast sensitivity function to represent contrast threshold as a function of presentation distance for a particular pixel scale. The result is presented in Figure 7.4, which indicates the minimum contrast between every other pixel required for the visual system to detect the corresponding spatial frequencies. In order to compromise between minimising the viewing distance to maximise the size of the text and maximising the viewing distance to prevent pixels from being resolved, a distance of 75 cm was chosen. This distance is indicated on Figure 7.4 and corresponds to a minimum contrast of 0.99 for the resolution of pixels. It is unlikely that we will present images with that have a contrast over 0.99 for such high spatial frequencies. If we wanted the pixels to be Nyquist sampled, such that there were two pixels per resolution element of the eye, we would need to present them at a spatial frequency of 6874 cycles per radian. This would correspond to a viewing distance of 1.7 m. At 0.75 m the maximum contrast should be 0.73. At a distance of 75 m, single letters subtended 15 arcmin, equivalent to an acuity of 20/60. Therefore a six letter target word subtended about 1.5 - 2.0 °, which is within the isoplanatic patch of a typical eye.

7.2.4. Amplitude of aberration

In order to keep to a reasonable experiment duration so that subjects did not become fatigued, we decided to test four amplitudes of each aberration. To choose a suitable range of amplitudes, we conducted a preliminary experiment in which we found that applying 0.3 - 0.5 μm rms defocus (over a 3.5 mm pupil) allowed subjects to read text, but with substantially increased difficulty. The main study used text samples aberrated with 0.3 μm, 0.35 μm or 0.4 μm rms of one of the three types of aberration. These values are measured over a 3.5 mm pupil and so correspond to 0.68 D, 0.79 D and 0.90 D of equivalent defocus. It should be noted that the text size used in this experiment was three times larger than the normal acuity limit (20/60 as opposed to 20/20) and so the impact of these aberrations

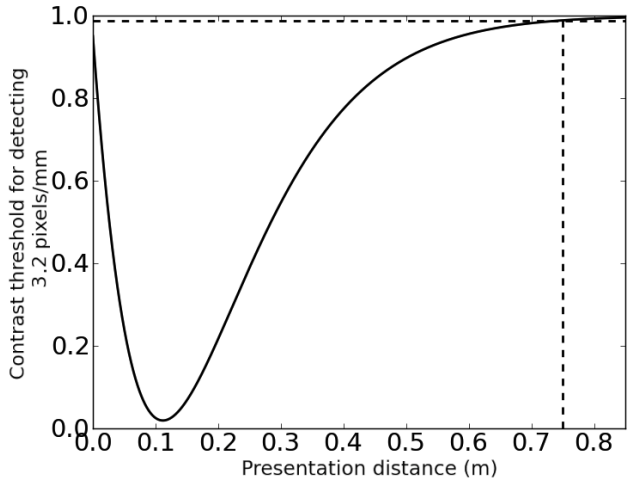

Figure. 7.4: Contrast threshold of the human visual system as a function of presentation distance for pixels on the monitor. The x-axis represents the distance at which 0.27 mm pixels correspond to spatial frequencies detectable by the visual system. A viewing distance of 0.75 m is indicated of the graph. The minimum contrast required for the visual system to detect this frequency is 0.99.

was lower than for a letter at the acuity limit.

In addition, a high amplitude was chosen for each type of aberration to find the result of the subject failing to read correctly. Since we expect the different types of aberration to effect reading by different amounts, the high amplitude used was specific to the aberration type. From the preliminary experiment we found that 0.55 μm of defocus caused significant reading difficulty and could be informally described as illegible. Equivalent amplitudes of coma and secondary astigmatism were determined by matching the area under their MTFs, which had been multiplied by the relative weights of spatial frequencies in the band used for letter recognition (Solomon & Pelli, 1994), to that of 0.55 μm of defocus. The equivalent amplitudes for coma and secondary astigmatism were 1.55 μm and 0.82 μm respectively. A control condition was also tested where no aberration was applied and results were compared to the control data. This avoided large errors caused by inter-subject variability in reading performance. Examples of the lowest amplitude for each aberration are given in Figure 7.1,

demonstrating the qualitative differences between the three types of aberration, at the same amplitude, when applied to text. Examples of each type and amplitude of aberration are given in Figure 7.5.

7.3. The Cambridge Research Systems high-speed video eyetracker

The text samples were presented on a CRT using a Cambridge Research Systems ViSaGe visual stimulus generator. The stimulus generator has a contrast resolution of 14-bits per gun, further details of this are given in Chapter 6. The average luminance of the monitor was 104 cd/m^2 (4002 Td) and the output of the display was gamma corrected. Eye movements were sampled at 250 Hz using a Cambridge Research Systems High-Speed Video Eyetracker (CRS HS-VET). This system has a spatial resolution of 0.05° and accuracy of 0.125°-0.25°. The CRT monitor and eye tracker were controlled using the CRS Matlab toolboxes.

Tracking technique and calibration

The CRS HS-VET measures the positions of the centre of the cornea and the the centre of the pupil. The centre of the cornea is determined by measuring the positions of two first Purkinje images (those created by reflections from the anterior surface of the cornea). The centre of the cornea and the centre of the pupil are offset when the optical axis of the eye is not perpendicular to the image plane. As the eye rotates, these centres move relative to each other in three dimensions, which can be modelled via:

$$\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \begin{bmatrix} C_x \\ C_y \\ C_z \end{bmatrix} + \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} + \begin{bmatrix} X_{offset} \\ Y_{offset} \end{bmatrix}, \tag{7.2}$$

Figure. 7.5: Examples of the stimuli used in the reading experiment. Images are scaled such that the are the size used in the experiment (15 arcmin) when held at a typical reading distance (23 cm).

where $T_{x,y,x}$ and $C_{x,y,z}$ are the coordinates of the target stimulus (the calibration spot) and measured offset of the pupil and first Purkinje image for a gaze position (projected on to the monitor in cartesian coordinates) given by X_0 and Y_0. The parameters a, b, c, d, e, f, g, h, α, β, γ, δ, X_{offset} and Y_{offset} are determined from calibration measurements made at known gaze coordinates (X_0 and Y_0) (Wallington & Symons, 2008). As a consequence of this matrix operation, errors in the y-coordinates can contaminate the x-coordinates. In this experiment eye movement was mostly in the x-direction, since sentences occupied a single line, but the accuracy of the eye movement data was constrained by errors in both the x and y direction. Figure 7.6 demonstrates the error on the measured eye position as a function of screen position, measured in a single subject with either a 9, 15 or 25 point calibration grid. This shows that the accuracy of the eye movement data is greatest at the centre of the monitor. This justifies placing the target word in the middle of the sentence since we need to know for certain which fixation corresponds to the target word. There is little difference between calibration conditions since these errors arise from non-linearities in the movement of the first Purkinje image and from centroiding an increasingly elliptical pupil, rather than from the calibration procedure itself. The calibration parameters will vary between subjects due to differences in individual eye anatomy but they will also vary during a session with the same subject due to changes in head position relative to the monitor. Calibration was therefore performed and verified every four trials and subjects were asked to remain still between calibration procedures. Calibration was performed using a 9 (3×3) point grid since fewer points would reduce the experiment duration.

7.4. Saccade detection algorithm

To calculate meaningful results from eye position data it is necessary to carefully remove blinks and noise so that eye movements can be detected. Noise removal is particularly important for velocity based detection methods since taking the derivative of the data will amplify any noise. Eye movement data can then be grouped into saccades and stationary

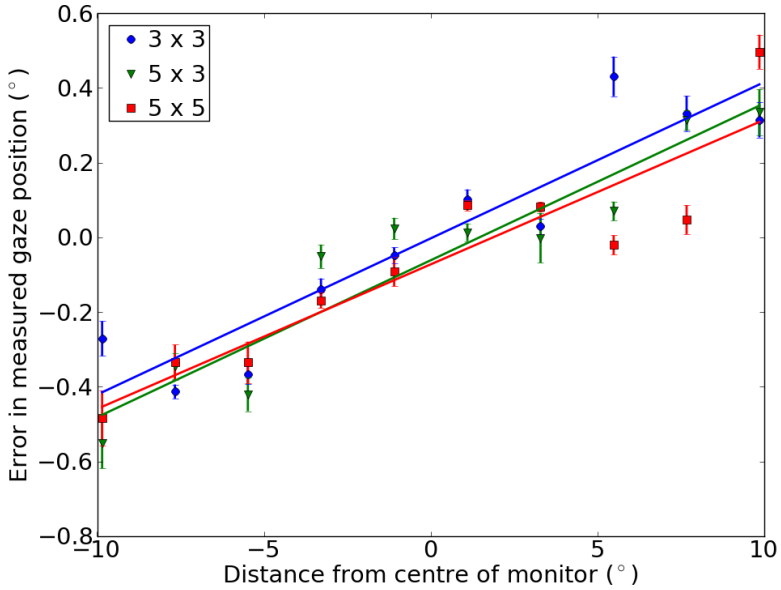

Figure. 7.6: Eye tracker measurement error as a function of gaze position, given as the visual angle. Data are given for 10 target fixation points spaced equally over the extent of the monitor. Targets were presented sequentially with adjacent targets only being displayed once fixation on the preceding target had been maintained for 750 ms. Data are averaged over 6 trials and the error bars represent the standard error on the mean. Three calibration conditions were tested with either 3×3, 5×3 or 5×5 points in the x and y directions respectively. These data show that the measurement error increases with visual angle and that there is little difference between calibration conditions. There is also a tendency for the variation in the measurement error to increase with visual angle, as shown by larger error bars away from the centre of the monitor.

periods in between can be considered to be fixations.

In this experiment eye movement data were analysed using an algorithm by Nyström and Holmqvist (2010) (article contains a copy of the code in Matlab, which was translated into Python). This algorithm was chosen as it demonstrates two main advantages over other algorithms. Firstly it has an adaptive velocity threshold for saccade detection that is calculated based on the noise. This removes the need for a user-defined estimate of threshold which can cause unreliable results, either detecting small noisy movements as saccades if the threshold is too small or neglecting slow saccades if it is too large. Secondly the algorithm identifies a separate class of eye movements, called glissades. These occur at the end of a saccade when the eye over-shoots its target landing position and corrects with a small saccade-like movement (Weber & Daroff, 1972; Kapoula, Robinson, & Hain, 1986). The presence of glissades in eye movement data depends on the method of eye tracking used and it is thought that they serve no useful purpose (Kapoula *et al.*, 1986). Classifying these eye movements separately prevents their presence, or absence, randomly affecting saccade and fixation data. The disadvantage of this algorithm, stated by Nyström and Holmqvist (2010), is that glissade-like movement preceding saccades are not detected (although this rarely occurs) and smooth pursuit eye movements are not separately defined. Smooth pursuit eye movements are unlikely to occur in reading since there are no moving targets for the eyes to track.

7.4.1. Smoothing and removing noise

In this algorithm noise is reduced using a Savitzky-Golay filter to smooth the data. Polynomials are fitted to the data in short windows using a finite impulse response filter that equates to a least squares fit. The impulse is a set of convoluting integers and there exists a set integers for any order derivative of the data. Using this technique the velocities and accelerations are determined analytically rather than numerically, which would amplify any noise in the data. This gives good performance for conserving high spatial frequency

information while maintaining local temporal and spatial information (Savitzky & Golay, 1964), which are often lost with moving average filters for example.

Noise in the data is initially determined using an indication from the eye tracker that tracking has been lost. This means that either the pupil centre or the corneal centre cannot be measured due to blinks, large eye movements which cause the first Purkinje images to be lost, or system errors. Data were also removed if the eye movement was not physiologically possible (velocity $> 1000°/s^2$ or acceleration $> 100000°/s^2$, see Duchowski, 2003, for example) or if the eye movement took fixation beyond the monitor. Additionally data points before confirmation that the subject was fixating at the start of the sentence were omitted.

7.4.2. Detecting saccades and fixations

Saccades are determined by finding velocity peaks above the saccade detection threshold. Once saccades are identified their onsets are found by searching backwards for the first sample to go below the saccade onset threshold. The saccade detection threshold and the saccade onset threshold are both determined from the global noise during a trial by iteratively adapting the estimate of the detection or onset threshold until the difference between successive estimates converges to within $1°/s$. The saccade offset is determined in the same way to the saccade onset, however it is calculated locally based on the noise within a 40 ms window preceding the saccade. This accounts for variation in the noise over the course of a single trial, for example due to the change in gaze position, as discussed in Section 7.3. Glissades are then detected by looking for velocity peaks above threshold that occur within 40 ms of the saccade offset. These are defined as either strong glissades, those which have a velocity peak above the saccade detection threshold, or weak glissades, those which have a velocity peak above the saccade onset threshold but not above the saccade detection threshold. The glissade onset is defined by the offset of the preceding saccade and the glissade offset is defined as the first sample at which the velocity changes direction. Everything that is not classified as noise, a saccade or a glissade is labelled as a fixation.

Examples of eye movement data taken during this experiment showing saccades, glissades and consequently fixations are given in Figures 7.7 and 7.8.

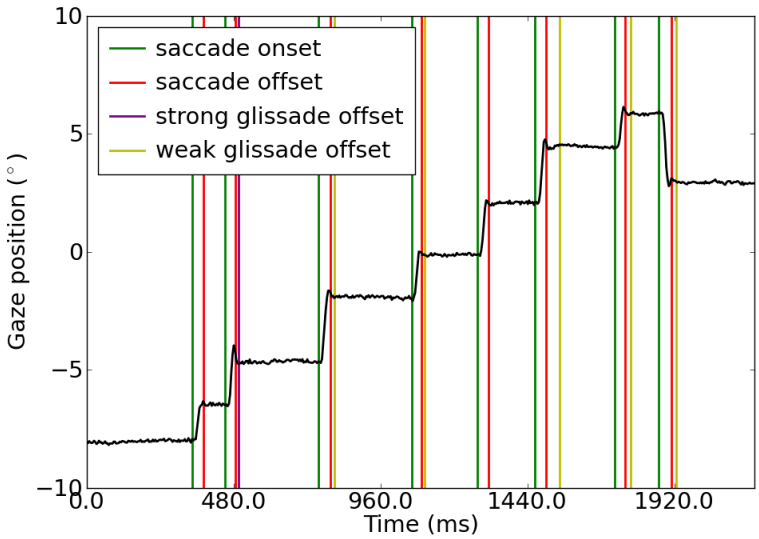

Figure. 7.7: Eye movement data while reading a sentence with no aberration. Gaze position refers to eye movements projected on to the horizontal axis of the monitor, expressed as visual angle. Saccade onset and offset and both strong and weak glissades are indicated by vertical lines. Strong glissades are those with a peak velocity above the saccade detection threshold and weak glissades are those with a peak velocity above the saccade onset threshold. Periods not corresponding to eye movements (either saccades or glissades) are counted as fixations. These fixations are shown in Figure 7.8.

Saccade-detection procedure

These sections have described the saccade detection algorithm (a copy of the code in Matlab can be found in Nyström & Holmqvist, 2010), which was implemented in python and followed these steps:

- Read in raw data from the eye tracker (point of gaze projected on to the screen in mm).
- Remove any data points that are considered to be noise (as described in Section 7.4.1).
- Smooth the data with a Savitzy-Golay filter that uses a window of 9 data points and fits

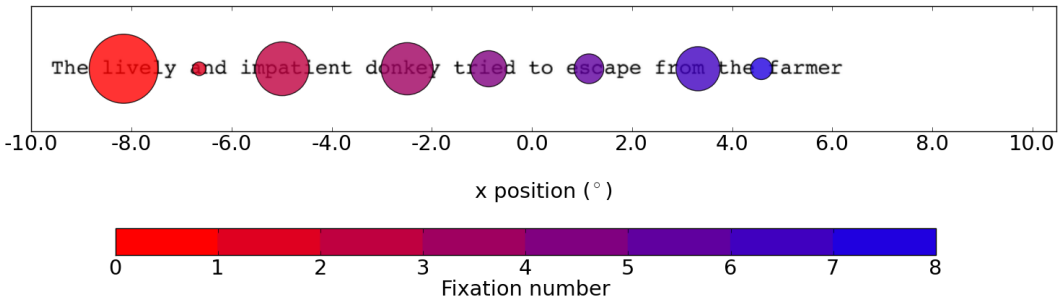

Figure. 7.8: Fixations made while reading a sentence with no aberration. The circles represent the location (position) and duration (diameter) of fixations. The shortest fixation was 72 ms and the longest fixation was 344 ms. The first fixation was not included in our analysis since it may partially represent time that the subject spent looking at the fixation cross. Fixations tend to land near to the middle of words, although the this is not completely represented here due to the measurement inaccuracies discussed in Section 7.3. These can cause fixation positions to be measured 1-2 letters from the true position. Short words such as "to" are skipped and longer words that provide context, such as "donkey" are fixated longer. The eye movement data corresponding to these fixations are given in Figure 7.7.

a polynomial of order 3. Using the analytical derivations performed by this algorithm calculate the position, velocity and acceleration of the gaze positions. Convert the smoothed data samples in to visual angle using the presentation distance.

- Calculated the saccade detection velocity threshold by initially setting the threshold estimate to the median velocity over the whole trial. The calculate the next threshold estimate as $\mu + 6\sigma$ (where μ is the mean and σ is the standard deviation) of the data below this threshold. Iterate until the difference between successive threshold estimates is less than 1°. The threshold at which this occurs is the saccade detection threshold.

- Calculate the saccade onset velocity threshold as $\mu + 3\sigma$ of the data below the saccade detection threshold.

- For each saccade (blocks of data that are above the saccade detection threshold) search backwards from the end for the first velocity sample to exceed the saccade onset threshold. This is the beginning of the saccade.

- Find the saccade offset by first finding the saccade offset velocity threshold which is $\mu + 6\sigma$ of the data in a 40 ms window immediately preceding the saccade. The search forwards from the beginning of the saccade until this velocity is found. This is the end of the saccade.

- Examine the samples following the end of the saccade for a glissade. Examine a 40 ms window following the end of the saccade for the last velocity sample that exceed the saccade detection threshold (indicating a strong glissade) or the saccade onset threshold (indicating a weak glissade). This defines the end of the glissade with the onset being the offset of the saccade.

- Label all data samples that are noise, or part of a saccade or glissade, as a fixation.

7.5. Experimental method

In this experiment subjects were presented with a single sentence with either a high or low lexical frequency target word embedded in the middle. Each subject was presented with 2 sentences per lexical frequency (high or low), type of aberration (defocus, coma or secondary astigmatism) and amplitude of aberration (0.3 μm, 0.35 μm, 0.4 μm or a high amplitude). This was also replicated for a control condition. This lead to 52 trials per subject during which their eye movements were monitored so that fixation times could be calculated.

7.5.1. Subjects

Nineteen subjects participated in this study, twelve male and seven female, with a mean age of 28 years (SD = 7 years). All subjects were fluent in English, had at least 17 years in education and normal or corrected-to-normal vision. Subjects read with natural pupil dilation using any vision correction (6 wore spectacles and 1 wore contact lenses) they would normally use for viewing a computer screen at a distance of 75cm.

7.5.2. Procedure

Subjects were positioned in a headrest and asked to read the sentences silently so as to understand them. Each sentence was followed by a comprehension question that required a yes-no response via a button box. As an example, if the sentence were "Sarah was trying to draw her sister but was finding it very difficult" the question would be"Did Sarah find the task easy?". The subject was required to press a button to start each trial so that they could control the pace of the experiment and take a break when required. Each sentence was preceded by a fixation cross that directed the subject's gaze to the location at which the start of the sentence would be presented. Each sentence was displayed and data collected only once it had been confirmed that the subject was fixating at the position of the first word. A 9-point calibration was performed and verified before every fourth trial or after the subject had moved from the headrest. Data were collected from subjects in single sessions of 45 minutes duration. Trials were excluded from the analysis if the subject did not correctly answer the comprehension question or the eye movement data contained more than 20% noise. Under these criteria, 9.3% of trials were excluded from the analysis. The accuracy of saccade detection was checked manually before any further analysis was carried out.

Pupil size was monitored during the experiment to check that the contribution of the subject's aberrations to the retinal image quality remained minimal. For each subject, the amplitude of higher-order aberrations was calculated over their average pupil diameter during the experiment. Over all subjects the average amplitude of higher-order aberrations was 0.018 ± 0.007 μm rms and the average pupil diameter was 3.5 ± 0.4 mm. Pupil diameters during the experiment were always smaller than the diameter over which this measurement was taken. This was accounted for by recalculating the Zernike coefficients over the smaller diameter. Histograms of subjects' higher-order aberrations are given in Figure 7.9 and 7.10.

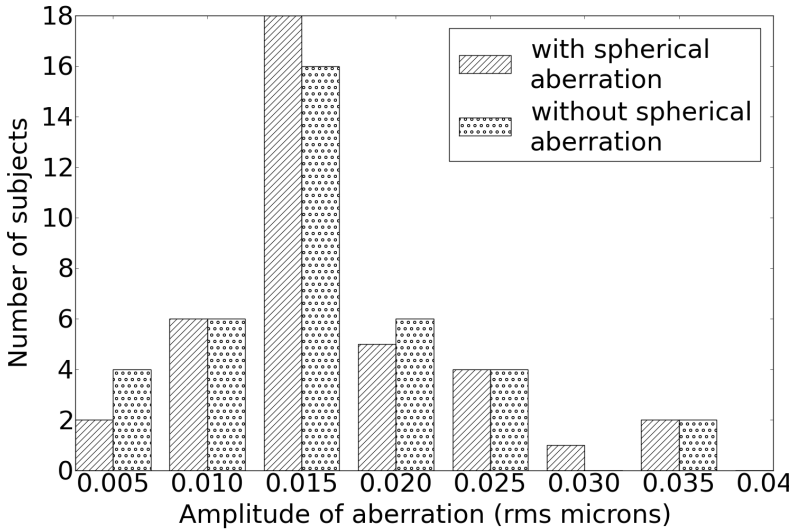

Figure. 7.9: Histogram of the rms value of subjects' measured higher-order aberrations over their average pupil diameter during the experiment. Only higher-order aberrations are included since subjects read using any prescribed vision correction, such as spectacles, which correct low order aberrations. Results are shown with and without the spherical component (Z_4^0) included as spherical aberration is affected by spectacle prescription.

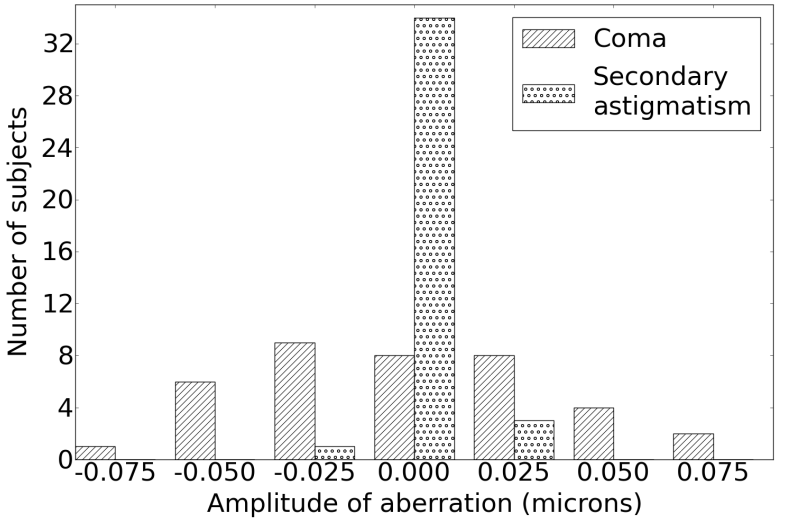

Figure. 7.10: Histogram of the rms value of subjects' measured coma (Z_3^1) and secondary astigmatism (Z_4^2) over their average pupil diameter during the experiment. For comparison the rms amplitudes of our simulated aberration were 0.3 μm, 0.35 μm and 0.4 μm.

7.6. Experimental results

For each trial the average fixation duration and the number of fixations were calculated as global measures of performance. For each participant the average for the four trials in the control condition was subtracted from these measures to avoid large errors from inter-subject variability. The weighted average, according to the number of valid trials contributing to the estimate, was then calculated over all subjects. Additionally, the sum of all fixation durations on the target word before leaving the word (gaze duration) and the sum of all fixations on the target word (total reading time) were calculated for each trial. These local measures were then used for the lexical frequency analysis. Data were subjected to repeated measures ANOVAs. When sphericity could not be assumed, a Greenhouse-Geisser correction was applied for epsilon values below 0.75 otherwise a Huynh-Feldt correction was used. In the reports below we quote corrected degrees of freedom.

7.6.1. Global measures

Figure 7.11 shows the increase in average fixation duration over the control condition vs. the amplitude of the aberration. The data were analyzed with a 3 (type of aberration) x 3 (amplitude of aberration) two-way repeated measures ANOVA. We found a significant main effect of the type of aberration ($F(2.0,143.3) = 59.9$, $MSE = 23177.4$, $p < 0.001$) with coma causing the smallest increase in average fixation duration (25.2 ms averaged over all amplitudes), followed by defocus (32.4 ms averaged over all amplitudes) and by secondary astigmatism which caused the greatest increase (46.3 ms averaged over all amplitudes). We also found a significant effect of the amplitude of the aberration ($F(2.0,143.3) = 15.7$, $MSE = 5840.6$, $p < 0.001$). As we expected there was an increase in the average fixation duration as the amplitude of aberration was increased with, on average, an extra 15.9 ms exhibited at 0.3 μm of aberration, 31.7 ms at 0.35 μm and 56.2 ms at 0.4 μm. Most importantly there was a significant interaction between type and amplitude of aberration ($F(3.6,143.3) = 9.7$, $MSE = 3299.5$, $p < 0.001$) indicating that the effect of increasing the amplitude

was different for different types of aberration. We explored the significant interaction with an analysis of simple effects. Since there were only three groups, where one-way ANOVAs revealed a significant effect, post-hoc pairwise comparisons were made via Fishers least significant difference (Howell, 1992, pp.356).

For defocus, there was a significant effect of the amplitude of aberration ($F(2.0,36.0) = 40.3$, $MSE = 11934.0$, $p < 0.001$) causing an increase in average fixation duration from 200.9 ms at 0.3 μm, to 214.1 ms at 0.35 μm, up to 250.6 ms at 0.4 μm. Pairwise comparisons identified differences between 0.30 μm and 0.35 μm ($t(18) = 3.0$, $p \leq 0.007$), between 0.35 μm and 0.40 μm ($t(18) = 7.8$, $p < 0.001$) and between 0.30 μm and 0.40 μm ($t(18) = 6.3$, $p < 0.001$).

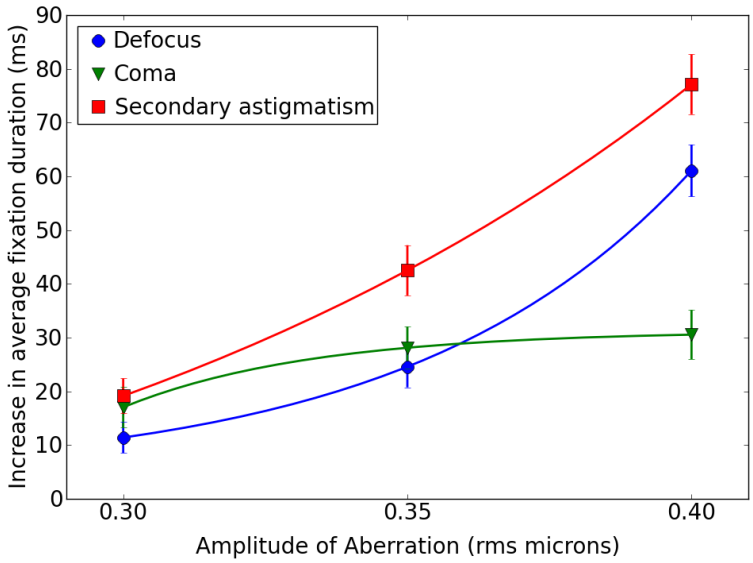

Figure. 7.11: Increase in the average fixation duration from control data over a whole sentence with increasing amplitude of aberration. Error bars represent the standard error on the mean of the measurements across 19 subjects. Three types of aberration are shown, defocus (Z_2^0), coma (Z_3^1) and secondary astigmatism (Z_4^2). Data are averaged over 19 subjects, each viewing 4 sentences per type and amplitude of aberration. Exponential growth curves have been fitted to the data via the least squares method. The growth constants were found to be 20.4 μm^{-1} for defocus (SSE < 0.001), -30.3 μm^{-1} for coma (SSE < 0.001) and 7.8 μm^{-1} for secondary astigmatism (SSE < 0.001)
.

For secondary astigmatism, there was also a similar significant effect of the amplitude of aberration ($F(2.0,36.0) = 33.0$, MSE = 16302.3, $p < 0.001$). Once again, the average fixation duration increased with larger aberrations from 208.6 ms at 0.3 μm, to 232.0 ms at 0.35 μm, up to 266.6 ms at 0.4 μm. Again, pairwise comparisons identified differences between 0.30 μm and 0.35 μm ($t(18) = 3.2$, $p \leq 0.005$), between 0.35 μm and 0.40 μm ($t(18) = 7.7$, $p < 0.001$) and between 0.30 μm and 0.40 μm ($t(18) = 5.0$, $p < 0.001$).

For coma, there was a significant, but smaller effect, of the amplitude of aberration for coma ($F(1.8,33.0) = 3.9$, MSE = 898.5, $p \leq 0.034$). The average fixation duration for coma increased from 206.5 ms at 0.3 μm, to 217.6 ms at 0.35 μm, up to 220.1 ms at 0.4 μm. Clearly these effects are much smaller in magnitude. In line with this, pairwise comparisons showed that the reliable effect was driven by differences between 0.30 μm and 0.35 μm ($t(18) = 2.6$, $p \leq 0.017$) and between 0.30 μm and 0.40 μm ($t(18) = 2.4$, $p \leq 0.028$) though not between 0.35 μm and 0.40 μm ().

These trends can be seen clearly in Figure 7.11, which shows that text becomes increasingly difficult to process as the amplitude is increased for defocus and secondary astigmatism. By contrast the difficulty of processing the text plateaus at 0.35 μm for coma. This pattern of effects is consistent with the suggestion that when viewing text with coma the structure of the letters is still distinguishable whereas with defocus and secondary astigmatism their structure is altered. This can be seen in Figure 7.1, which shows that coma creates a smearing effect across the letters but leaves their form relatively intact.

Number of fixations

Figure 7.12 shows the increase in the number of fixations over the control condition vs. the amplitude of the aberration. Another two-way repeated measures ANOVA showed a significant main effect of the type of aberration ($F(1.7,94.5) = 8.5$, MSE = 60.7, $p \leq 0.002$), with coma again causing the smallest detriment (no extra fixations on average over

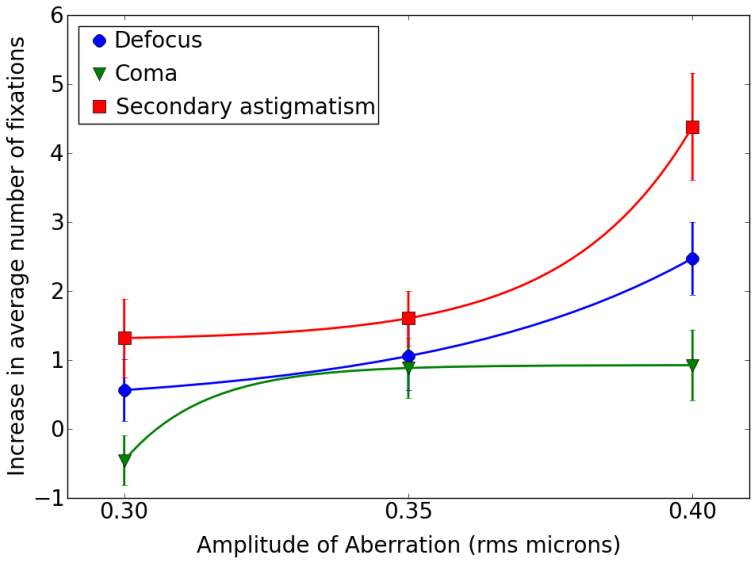

Figure. 7.12: Same format as Figure 7.11. Increase in the average number of fixations from control data over a whole sentence with increasing amplitude of aberration. Exponential growth curves have been fitted to the data via the least squares method and the growth constants were found to be 21.1 μm^{-1}for defocus (SSE $<$ 0.001), $-$70.1 μm^{-1} for coma (SSE $<$ 0.001) and 45.3 μm^{-1}for secondary astigmatism (SSE $<$ 0.001).

all amplitudes), followed by defocus (1 extra fixation on average over all amplitudes) and by secondary astigmatism which caused the largest increase in the number of fixations (2 extra on average over all amplitudes). We also found a significant effect of the amplitude of the aberration ($F(1.3,94.5) = 12.3$, MSE = 88.4, $p \leq 0.001$). Consistent with what we expected, there was an increase in the number of fixations and this grew from 0 at 0.3 μm, to 1 at 0.35 μm, up to 3 at 0.4 μm.

Since there were only three groups we explored these main effects with pairwise comparisons made via Fishers least significant difference, which revealed that the main effects were driven by the following differences: There was a significant difference in the number of fixations between secondary astigmatism and defocus ($t(18) = 1.9$, $p \leq 0.008$) and between secondary astigmatism and coma ($t(18) = 3.5$, $p \leq 0.002$). There was only a marginal difference between defocus and coma ($t(18) = 3.0$, $p = 0.078$). The data indicate that

defocus was least disruptive, causing readers to make fewest fixations, coma slightly more so, with secondary astigmatism most disruptive. There was also a significant difference in the number of fixations between 0.30 μm and 0.35 μm ($t(18) = 2.6$, $p \leq 0.018$), between 0.35 μm and 0.40 μm ($t(18) = 3.4$, $p \leq 0.004$) and between 0.30 μm and 0.40 μm ($t(18) = 3.8$, $p \leq 0.001$). This shows that the number of fixations increased significantly across all amplitudes of aberration. Although there was no significant interaction ($F(2.1,94.5) = 2.5$, MSE = 17.1, $p \leq 0.091$) the data show similar numerical trends to those shown in Figure 7.11.

To briefly summarise, these global measures indicate that the amplitude of aberration reliably influenced both the number and duration of fixations during reading. We also saw a significant influence of the type of aberration on both of these measures, with poorest performance for secondary astigmatism and best performance for coma. For fixation duration we additionally note that increasing the amplitude was differentially effective in impairing performance with different types of aberration.

7.6.2. Local measures

The gaze durations and total reading times for high and low frequency words are given in Figures 7.13 and 7.15. The differences in gaze durations and total reading times between the high and low frequency target words (i.e. the frequency effect) are given in Figures 7.14 and 7.16.

Average gaze durations on target word

The average gaze durations were subjected to a 3 (type of aberration) x 3 (amplitude of aberration) x 2 (lexical frequency) three-way repeated measures ANOVA. As expected there was a significant main effect of the amplitude of aberration ($F(1.6,378.4) = 33.5$, MSE = 153954.2, $p < 0.001$) with durations increasing on average from 223.6 ms at 0.3 μm,

to 251.4 ms at 0.35 μm, up to 294.3 ms at 0.4 μm. This pattern mirrors that reported earlier. We also found a significant effect of the lexical frequency of the word ($F(1.0,378.4) = 20.9$, MSE = 112180.1, $p < 0.001$), as expected, with fixation durations if 238.2ms for high frequency words and 274.7ms for low frequency words. There was a marginal effect of the type of aberration ($F(1.9,378.4) = 3.2$, MSE = 46022.3, $p \leq 0.057$), with subjects fixating the longest in the presence of secondary astigmatism (279.5 ms on average), followed by defocus (256.4 ms on average) and the shortest fixations were those in the presence of coma (233.4 ms on average). A marginal interaction between the type and amplitude of the aberration was also found ($F(2.3,378.4) = 2.9$, MSE = 29404.2, $p \leq 0.063$) indicating that, as in the global results, the increase in fixation duration was different for the different types of aberration. None of the other two-way interactions, nor the three-way interaction were significant.

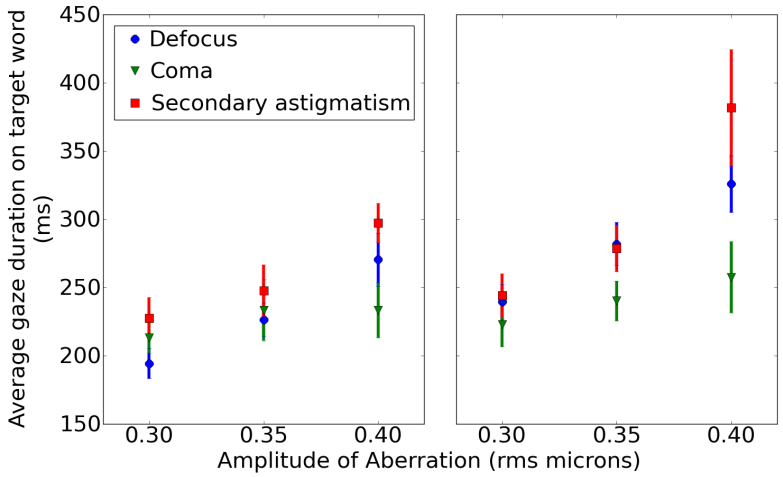

Figure. 7.13: Average gaze durations for a) high and b) low lexical frequency target words. Error bars represent the standard error on the mean of the measurements across 19 subjects.

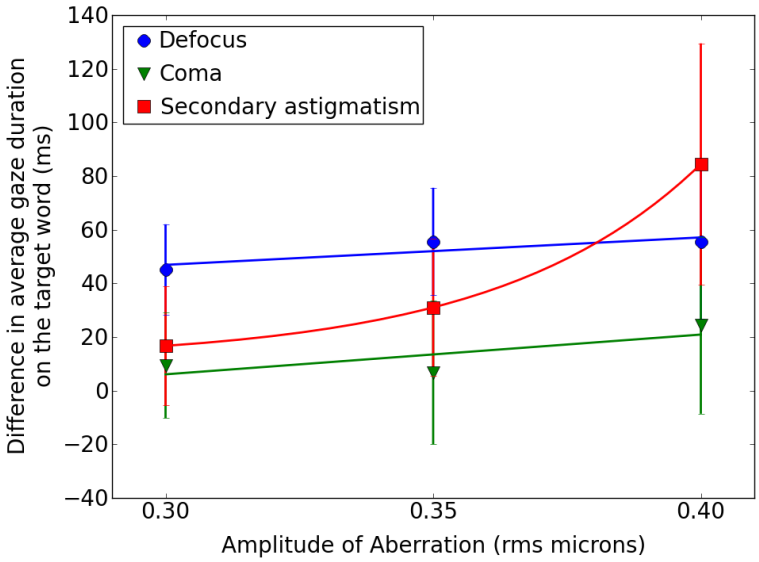

Figure. 7.14: Difference in the average gaze durations between the low and high lexical frequency words (low minus high). Error bars represent the standard error on the mean of the measurements across 19 subjects. These error bars are larger than in Figure 7.11 as these data are not calculated with respect to the control condition and so include inter-subject variability. Exponential growth curves have been fitted to the data via the least squares method and the growth constants were found to be 0.003 $\mu m^{-1}$for defocus (SSE = 19.2), -0.002 $\mu m^{-1}$for coma (SSE = 69.0) and 26.3 $\mu m^{-1}$for secondary astigmatism (SSE < 0.001).

Total reading times on target word

The total reading times were also subjected to a 3 (type of aberration) x 3 (rms amplitude of aberration) x 2 (lexical frequency) three-way repeated measures ANOVA. In line with what we expected there was a significant main effect of the amplitude of aberration ($F(1.6,233.8) = 10.9$, MSE = 1531859.0, $p \leq 0.005$), with durations increasing on average from 380.7 ms at 0.3 μm, to 477.0 ms at 0.35 μm, up to 615.8 ms at 0.4 μm. Consistent with our other results, there was also a significant effect of the type of aberration ($F(1.6,233.8) = 6.5$, MSE = 1055920.3, $p \leq 0.002$), with subjects fixating the longest in the presence of secondary astigmatism (602.4 ms on average), followed by defocus (493.9 ms on average) and the shortest fixations were those in the presence of coma (377.2 ms on average). Unsurprisingly, we also found a significant effect of the lexical frequency of the word ($F(1.0,233.8)$

= 1.7, MSE = 134175.6, p $\leq$ 0.001) with subjects fixating on high frequency words for 427.0 ms on average and on low frequency words for 555.4 ms. Importantly there was there was a significant interaction between the type of the aberration and the lexical frequency of the word (F(1.4,233.8) = 15.2, MSE = 2024411.0, p $\leq$ 0.040). There was also a significant interaction between the amplitude of the aberration and the lexical frequency of the word (F(1.9,233.8) = 1.2, MSE = 127891.5, p $\leq$ 0.009). In agreement with our previous findings, a marginal interaction between the type and amplitude of the aberration was also found (F(1.7,233.8) = 3.9, MSE = 637574.6, p $\leq$ 0.054). The three-way interaction was not significant.

On a priori grounds we decided to investigate the interaction between the type and

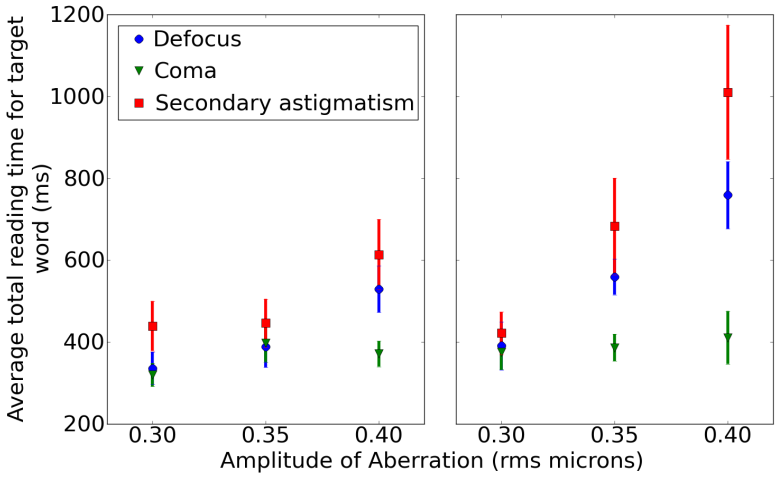

Figure. 7.15: Average total reading times for a) high and b) low lexical frequency target words. Error bars represent the standard error on the mean of the measurements across 19 subjects.

the amplitude of aberrations for high and low lexical frequency words separately. A 3 x 3 two-way repeated measures ANOVA showed no significant effect of this interaction on either the average gaze durations or the total reading times of high frequency words. There was however a significant effect of this interaction on the total reading times of low frequency words (F(2.1,79.8) = 3.4, MSE = 231524, p $\leq$ 0.038), indicating that the increase in total

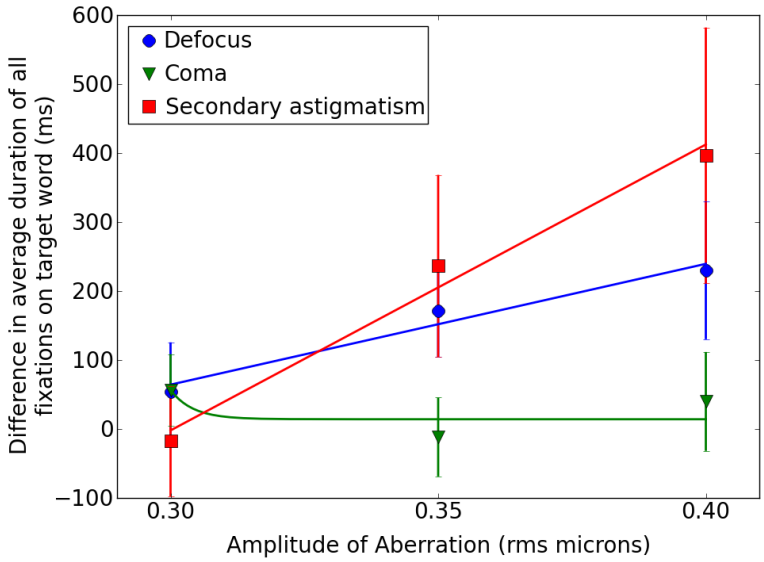

Figure. 7.16: Difference in the average total reading times between the low and high lexical frequency words (low minus high). Error bars represent the same value as for Figure 7.14. Exponential growth curves have been fitted to the data via the least squares method and the growth constants were found to be 0.004 $\mu m^{-1}$for defocus (SSE = 562.4), -241.4 μm^{-1} for coma (SSE = 1327.0) and 0.004 $\mu m^{-1}$for secondary astigmatism (SSE = 1474.6).

reading time of low frequency words was different for the different types of aberration. This interaction did not reach significance for average gaze durations (F(3.4,108.9) = 2.0, MSE = 16039, p = 0.118).

We investigated the significant interaction with an analysis of simple effects. Where one-way ANOVAs revealed a significant effect, pairwise comparisons were again made via Fishers least significant difference since there were only three groups. For defocus, there was a significant effect of the amplitude of aberration on the total reading times of low frequency words (F(1.9,34.2) = 8.8, MSE = 565732, $p \leq 0.001$). In the presence of defocus the total reading times of low frequency words increased from 390.8 ms at 0.3 μm, to 559.7 at 0.35 μm, up to 758.9 at 0.4 μm. Pairwise comparisons showed differences between 0.30 μm and 0.35 μm (t(18) = 2.4, $p \leq 0.028$), between 0.35 μm and 0.40 μm (t(18) = 2.3, $p < 0.037$) and between 0.30 μm and 0.40 μm (t(18) = 3.8, $p < 0.001$).

For secondary astigmatism, there was a significant effect of the amplitude of aberration ($F(1.4,25.2) = 6.3$, MSE = 2841004, $p \leq 0.012$). In the presence of secondary astigmatism the total reading times of low frequency words increased from 421.6 ms at 0.3 μm, to 682.8 at 0.35 μm, up to 1010.5 at 0.4 μm. Pairwise comparisons revealed differences between 0.30 μm and 0.35 μm ($t(18) = 2.3$, $p \leq 0.036$), between 0.35 μm and 0.40 μm ($t(18) = 2.3$, $p < 0.037$) and between 0.30 μm and 0.40 μm ($t(18) = 3.8$, $p < 0.001$). For coma, the one-way ANOVA was not significant and so no pairwise comparisons were made.

To summarise, the low frequency words took longer to identify than high frequency words and, when increasing amplitudes of defocus or secondary astigmatism were applied to the text, lexical identification became increasingly difficult. When text was viewed with coma aberration lexical identification was not similarly increasingly inhibited with increased amplitude of aberration. Once again, we consider that this indicates that coma leaves the form of letters and words relatively intact.

These local results with respect to the type and amplitude of the aberration show the same pattern as our global results for the average fixation duration but with an exaggerated effect on low, compared with high frequency words. The main effects of lexical frequency indicate clearly that lexical identification occurred even when readers were viewing aberrated text. This reinforces our conclusion based on the comprehension data that readers successfully processed the text. Furthermore significant two-way interactions involving lexical frequency for total reading time suggest that different types of aberration differentially affect lexical identification and high amplitudes of aberrations have a particularly strong effect. Differences between average gaze durations and total reading times are interesting because they reveal differences in behavior. An increase in average gaze durations would suggest subjects made additional refixations on the target word before moving on to the next word, whereas an increase in the total reading times would indicate that subjects revisited the target word. The absence of two-way interactions involving lexical frequency for gaze duration suggests that the size of the lexical frequency effect is affected by the type

and amplitude of the aberration in terms of later refixations on the word, after the first pass through the sentence.

7.6.3. The high amplitude conditions

The size of the lexical frequency effect in the high amplitude conditions showed that subjects were not likely to be identifying the words in the sentences. For defocus and coma the lexical frequency effect was inverted with high frequency words being fixated for 4 ms and 19 ms longer than low frequency words. For secondary astigmatism the lexical frequency effect was not inverted, however it's magnitude was reduced to 34 ms, which was lower than in the 0.4μm condition. This combined with the significantly reduced percentage of correct responses to comprehension questions (as shown in Figure 7.17) indicates that subject were not likely to be identifying the words in the sentences. In this high amplitude condition, increase in average fixation durations over a whole sentence (compared to the control condition) grew to 166 ms for defocus, 230 ms for coma and 195 ms for secondary astigmatism. These figures indicate the additional time subjects will spend, on average, looking at a word before giving up and moving on to the next word. This duration was smallest for defocus, although comprehension was better in this condition. This duration was larger for coma than secondary astigmatism, perhaps indicating that secondary astigmatism was more difficult to read as subjects gave up sooner. The increase in the number of fixations during a sentence (compared to the control condition) grew to 13 for defocus, 16 for coma and 11 for secondary astigmatism, compared to at most 4 fixations in the lower amplitude conditions. This indicated that subjects were finding the sentences significantly harder to read and these figures perhaps suggest that subjects will make an extra 11-16 fixations on average before giving up.

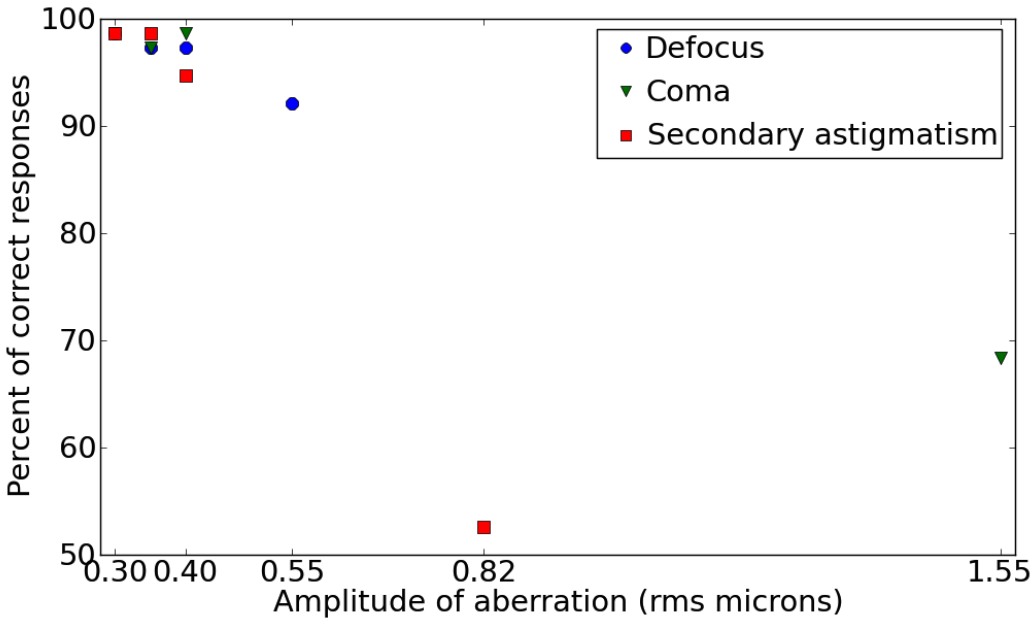

Figure. 7.17: Percentage of correct responses for each type and amplitude of aberration. For the high amplitude condition subjects responded correctly to 92.1% of the comprehension questions for defocus, 68.4% for coma and 52.6% for secondary astigmatism. For the lower amplitudes subjects correctly responded to at least 95% of the comprehension questions.

7.6.4. Comparison with confusion analysis

We compared the confusion metric for the three Zernike modes at the amplitudes and text size used in the experiment with the measured increase in average (global) fixation duration. Figures 7.18 and 7.19 shows there was a correlation between the predicted and measured performances that followed an exponential growth. The growth constant of the curve depended on the Zernike mode with coma having the lowest constant (-296.4 μm^{-1} in Figure 7.18 and -193.6 μm^{-1} in Figure 7.19). Defocus and secondary astigmatism had similar growth constants and we found that both sets of data could be described by the same curve, which had a growth constant of (198.8 μm^{-1} in Figure 7.18 and 56.7 μm^{-1} in Figure 7.19). This suggests that the mechanism by which secondary astigmatism and defocus affect performance might be similar. Coma however showed a larger increase in average fixation time for a smaller confusion metric suggesting performance is affected differently.

Again, we suggest that coma leaves the forms of letters relatively intact and so increases their confusability to a much lesser extent than defocus or secondary astigmatism.

Comparison with letter experiment

The amplitudes of aberration used in this experiment were different to those used in the letter experiment. This was because subjects used a different pupil size and the letters subtended a different angular size. We compared the amplitudes of aberration using the VSOTF (see Chapter 2) to find the amplitudes that would give similar image qualities in the letter experiment to those in the reading experiment. The equivalent amplitudes for defocus, coma and secondary astigmatism were approximately the same: 0.57 μm (equivalent to 0.3 μm), 0.67 μm (equivalent to 0.35 μm) and 0.77 μm (equivalent to 0.4 μm). These values overlap with the amplitudes used in the letter experiment (0.5 μm, 0.6 μm, 0.7 μm, 0.8 μm and 0.9 μm), therefore the images qualities were similar in each experiment. The confusion values were also similar, as can be seen from Figures 7.18 and 6.5 for example. In comparison to the letter recognition experiment these results different in an important way. In the letter experiment all three types of aberration could be predicted using the same relationship between the increase in contrast threshold and confusability. However, in comparison to the increase in average fixation duration defocus and secondary astigmatism can be predicted by the same relationship but coma cannot. The lower confusability values for coma indicate that, as we have already suggested, the forms of letters are less affected. The increase in average fixation duration is in fact higher than would be predicted by the black line in either Figure 7.18 or Figure 7.19. We suggest that this result bolsters our hypothesis that the effect of coma is more strongly associated with reading-specific performance degradations where letters are not presented in isolation. These effects could be due to lateral masking between words and letters within words (Townsend, Taylor, & Brown, 1971). Blurring of the letters causes their spatial extent increase and potentially overlap. If the spatial extent of a letter overlaps with an adjacent letter there is a reduction in contrast of that

adjacent letter. This lowers the signal to noise ratio of the adjacent letter, which impedes efficient feature and letter identification. Even if the spatial extent of a letter does not physically overlap with the adjacent letter there are likely to be crowding effects, which are caused by its close proximity. This interferes with the efficient integration of signals in the visual system such that information from adjacent letters is cross-contaminated. This noise may cause the feature and letter identification stages to be impaired resulting in a slower convergence on the correct identity of letter. Letter recognition is slower, though no less accurate and so word identification continues unhindered once the identity of the letter has been determined. This is reflected in the lack of an effect of the amplitude of coma on the size of the lexical frequency effect. Increasing amplitudes of coma did not increasing impair word identification. It is also possible that coma caused a stronger effect on saccade planning, which is not affected by the current word under fixation.

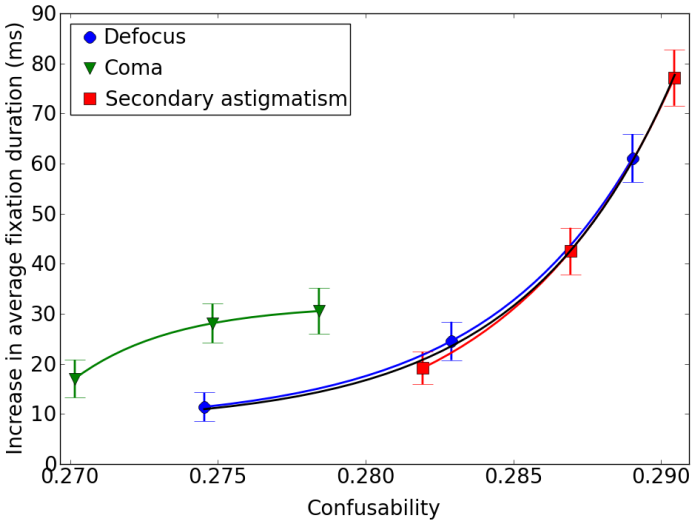

Figure. 7.18: Comparison between confusion analysis (comparing aberrated and aberrated letters) and the increase in average fixation time during reading, as shown in Figure 7.11. The mean fixation duration in the control condition was 189.5 ± 9.8 ms. Error bars represent the standard error on the mean of the measurements across 19 subjects. Exponential curves have been fitted to the data using the least squares method. The growth constants are 190.1 $\mu\text{m}^{-1}$for defocus (SSE < 0.001), -296.385 μm^{-1}or coma (SSE < 0.001) and 179.1 $\mu\text{m}^{-1}$for secondary astigmatism (SSE < 0.001). The data for secondary astigmatism and defocus can also be fit with the same curve with a growth constant of 198.8 μm^{-1}(SSE $= 4.535$), as shown in black.

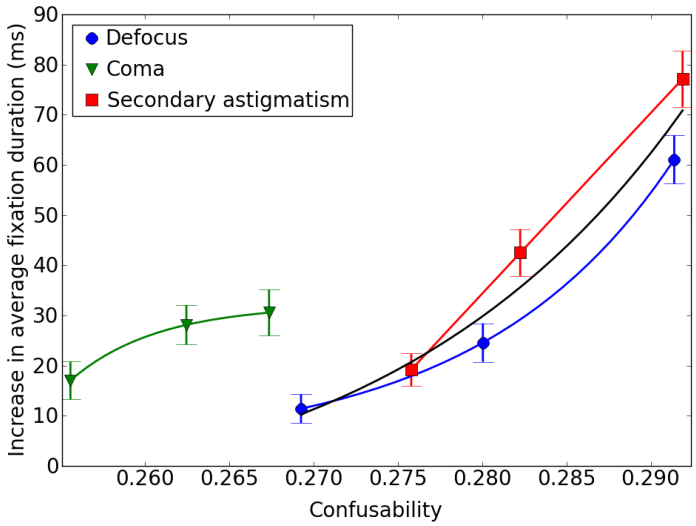

Figure. 7.19: As for Figure 7.18 but with the confusion compared between aberrated and unaberrated letters. The growth constants are 87.2 μm^{-1}for defocus (SSE $<$ 0.001), $-$193.641 μm^{-1}or coma (SSE $<$ 0.001) and 0.024 μm^{-1}for secondary astigmatism (SSE = 0.006). The data for secondary astigmatism and defocus has been fit with the same curve with a growth constant of 56.7 μm^{-1}(SSE = 174.553), as shown in black.

7.7. Discussion

Although Zernike coefficients are the standard quantity for measuring aberrations in the eye, their rms amplitude is not necessarily the most suitable metric for describing visual performance. This is because the effect of a rms amplitude of an aberration on visual performance is dependent on the Zernike mode in question. Understanding how visual performance is affected by higher-order aberrations is important when considering a correction for them. We wanted to know which aberrations contribute to a degradation in performance in a high-level visual task that includes both perceptual and cognitive components. Two factors that can affect fixation durations on a word during reading are the ease of word identification and uncertainty about where to target the next saccade. As pointed out in Chapter 3, word identification is a primary determinant of when a reader moves their eyes during reading (Liversedge & Findlay, 2000; Rayner, 1998). Degradation of the orthographic input slows word identification and other subsequent language processing stages. Saccadic targeting can also be affected by the visual clarity of the target, in this case the next word in the sentence to be fixated, and slower initiation of a saccade results in an increased fixation duration. The global analysis of average fixation duration subsumes both these factors. The lexical frequency manipulation allows us to isolate effects associated with word identification. Optical distortions, like those we have simulated, degrade visual information in a number of ways. Most simply they cause a reduction in contrast, however, they additionally cause changes in form by, for example, introducing new contours. They also increase the spatial extent of a stimulus and, in the case of text, they cause the letter and word boundaries to become less clear. The effect of coma could be informally described as smearing the text, leaving the form of letters relatively undisturbed but filling in the spaces between words and between letters within words, effectively increasing lateral masking. This characteristic of coma leads to a relatively low value of our letter-based confusion metric for a given rms amplitude. Defocus and secondary astigmatism have different characteristics. Like coma they increase the spatial extent of stimuli but in addition to this they have a profound effect on spatial form, as quantified by a higher confusion metric for these aberrations. In

interpreting our results we take these differences into consideration.

From the global results presented here we can see that the effect of higher-order aberrations on reading performance depends on both the type and the amplitude of the aberration. The most striking result is that coma does not affect reading performance as much as defocus or secondary astigmatism. This is in contrast to the findings of Oshika *et al.* (2006) who concluded that coma had a significant impact on letter contrast sensitivity. Applegate *et al.* (2002) and Applegate, Marsack, *et al.* (2003) found that Zernike modes with low angular order affect visual acuity more than those with high angular order, although in general it is the modes at the extreme edge of the Zernike pyramid that are less detrimental compared to those that are closer to the centre. Coma has a lower angular order than secondary astigmatism but we have not tested any modes from the extreme edge of the pyramid to conclusively agree or disagree with this finding. We also found that secondary astigmatism had a greater effect on the average fixation duration than defocus for the same rms amplitude, which agrees with the work of Fang *et al.* (2009) who found that fourth order aberrations affected optical quality the most. This finding suggests that it is as important (if not more so for improving reading performance) to consider correcting secondary astigmatism as it is defocus if the rms amplitude is high. Since our results differ from those of Applegate *et al.* and Oshika *et al.*, we suggest that the impact of higher-order aberrations on visual performance is likely to be dependent on the visual task and measure of performance.

We compared our results to our prediction of performance based on how similar letters are made to look in the presence of an aberration. We found a correlation between our measure of confusability and the increase in average fixation duration that could be described by an exponential curve. It is interesting to note that the relationships for defocus and secondary astigmatism are similar whereas the relationship for coma is clearly different. We believe this indicates a different source of performance loss for coma than for defocus and secondary astigmatism. For a given amplitude of aberration, the letters are more distinguishable for coma and the average fixation durations are shorter. However, in comparison

to defocus and secondary astigmatism, the average fixation durations are longer for lower confusability values. Given the qualitative difference between the distortions produced by different aberrations we suggest that performance loss in the case of coma might be more heavily influenced by saccade planning and lateral masking than by word identification (as we discuss below with respect to lexical frequency effects). Our data do not allow us to test this interpretation directly. With the current number of position measurements per target word, we were unable to reliably compare the changes in fixation location relative to the changes in the centres of gravity of the stimuli.

Lexical frequency effects can be taken as a direct index of the ease with which a word is identified. By considering how lexical frequency effects change between different types and amplitudes of aberration, we can assess the impact of different aberrations on word identification per se. An increased lexical frequency effect under conditions of aberration implies that low frequency words (those that are less common and therefore most difficult to identify) became even more difficult to identify when their constituent orthography was degraded. At the level of individual letters, this relates directly to the confusion analysis we have performed. Orthographic familiarity affects the ease of word identification which, as previously stated, contributes significantly to the fixation duration on the current word.

The responses to the comprehension questions confirm that the increases in average fixation duration are not an artefact of the text being made completely illegible. Clearly participants fully understood the sentences. The eye movement data additionally show that a lexical frequency effect was observed even in the presence of an aberration, indicating that subjects did successfully identify the words in the sentences (even if they had to revisit the words in order to do this). The interactions between either the type of aberration or the amplitude of the aberration and the lexical frequency of the word were significant for the total reading time but not for the gaze duration. In other words, the size of the lexical frequency effect was modulated by the type and amplitude of aberration only in relation to second pass fixations on the word, not the first pass fixations. For coma, the size of the

lexical frequency effect was constant across all amplitudes of aberration, for both gaze durations and total reading times. This suggests that subjects successfully lexically identified target words in the presence of coma during the first pass reading and did not need to spend a substantial amount of time refixating on them in order to identify them. For defocus and secondary astigmatism the lexical frequency effect was unaffected by the amplitude of aberration during the first pass through the sentence. However, there was an increase in the size of the lexical frequency effect with increasing amplitude of these aberrations that occurred in the refixations on the target word. We interpret this as a failure to initially correctly identify the words of the sentence during the first visit. Presumably, subjects either preliminarily guessed the identity of the target word or instead left the target without settling on its identity, in order to fixate words downstream in the sentence to (potentially) recruit further linguistic information that could facilitate it's identification. Either way, the data suggest that subjects made refixations on the target word to try and extract more orthographic information to either confirm, or to unambiguously identify them. Clearly, the visual information obtained on the first attempt was insufficient for word identification. The total reading times show that the lexical frequency effect was differentially modulated by the nature and level of aberration. An increase in the total reading times of low frequency words shows that when subjects were presented with words that are difficult to identify, the addition of an aberration increases this difficulty. This increased with the amplitude of aberration for defocus and secondary astigmatism indicating that these aberrations had a particularly strong effect on this psychological subprocess.

As already suggested in the discussion of global results, we predicted that defocus and secondary astigmatism should affect letter distinguishability more than coma. Global results show that the relationship between reading impairment and letter confusability is the same for defocus and secondary astigmatism. We suggest that our local results reinforce this interpretation since they show defocus and secondary astigmatism have an increasing impact on word identification, whereas coma does not. The confusion analysis has shown that it is

likely that coma causes a different reading behaviour to defocus and secondary astigmatism but not a different letter identification behaviour. We therefore suggest that this result is specific to reading words or letter strings and that the effect of some aberrations may be task-dependent, even when those tasks are related. We tentatively conclude that coma has a stronger effect on lateral masking and possibly saccade planning whereas defocus and secondary astigmatism have a greater impact on linguistic processing. We speculate that the reading strategy adopted when viewing text that had been distorted with secondary astigmatism, and to a lesser extent defocus, results from spatial phase changes. These result in spurious resolution, and the creation of sharp but false features. These false features provide incorrect orthographic cues that cause impaired lexical identification. Recent work by Ravikumar *et al.* (2010) has shown that phase errors caused by higher-order aberrations have an effect on visual acuity and that this effect is small for coma.

We found that on average subjects made 11-16 extra fixations when reading text that is so degraded that they could not identify words. These fixations were around 200 ms longer than when reading clear text. Subjects appeared to make more and longer fixations when reading text with coma in the case where they could not identify the words. We suggest that this is a consequence of subjects persisting in their attempt to read longer for text that appears less distorted and giving up sooner on text which contains letters which are highly confusable and difficult to identify. Despite this, comprehension was still poor for coma.

7.8. Conclusions

We conclude that during a task such as reading, there is a significant impact from the addition of higher-order aberrations that depends on both the type and magnitude of the aberration. We found the greatest impairment to reading was caused by secondary astigmatism and the smallest impairment by coma. The addition of an aberration results in less efficient extraction of orthographic information from the image, which in turn hinders

efficient lexical identification, and this encumbrance was greater for less familiar than more familiar words.

Importantly, there was a difference between the effects of defocus and secondary astigmatism and those of coma. We suggest that this arises due to the nature of the reading task, which is different from single letter identification. Defocus and secondary astigmatism had a clear impact on word identification, whereas we believe the effect of coma on reading is more strongly attributed to lateral masking effects and possibly disruptions to saccade planning. It is clearly important to consider differences between visual tasks when examining the effect of an aberration on visual performance.

Chapter 8

Conclusion

8.1. Summary

This thesis is concerned with the measurement of reading performance as a function of the type and amplitude of an aberration applied in the rendering of the text. In particular it may be valuable to know which types of aberration are worth correcting and which types one should avoid introducing through procedures such as refractive surgery. An example of such procedures is laser-assisted in-situ keratomileusis (LASIK) surgery in which the corneal surface is ablated with a laser to change its curvature and hence focal power. This process can introduce aberrations as well as correct them (see Oshika, Klyce, Applegate, Howland, & El Danasoury, 1999, for example). We have discussed the measurements of ocular aberrations and the effects of three types of aberration on visual performance. We limited our experiment to three types of aberration only to keep the duration of the experiment to a comfortable level for our subjects. However, other types and combinations of types of aberration are important. These three aberrations, defocus, coma and secondary astigmatism, were chosen using a prediction of their effects on letters, based on cross-correlation. We have compared our results with this prediction in order to infer the mechanisms that cause performance loss. Here we present a summary of the main findings of this thesis.

Rendering aberrations

We started this investigation by considering the differences between applying an aberration in the rendering of the stimuli and applying it optically, with an AO system for example. We can summarise this account with the simple conclusion that, when aberrations are imposed in rendering, it is important to measure the aberrations in subjects' eyes and make all possible attempts to minimise the effect of these aberrations on the retinal image. In our experiments, this was achieved by minimising the pupil diameter either artificially with an aperture or naturally by setting a high mean luminance value for the monitor. Accommodative error was minimised by providing high contrast cues, such as a fixation cross or box, to drive appropriate focus. In doing this we aimed to make the retinal image a true representation of the effect an induced aberration had on a stimulus and to make this representation consistent across subjects.

Repeatability of ocular wavefront measurements

We discussed the measurement of ocular aberrations and characterised the device we used in terms of its repeatability. We began by measuring the intrinsic repeatability errors of the system by using a model eye. These results showed that the standard deviation of repeated measurements was similar across all Zernike coefficients that were calculated (up to and including fifth radial order). However, there were large coefficients of variation ($\frac{SD}{mean}$) for some modes indicating that, while the absolute error (SD) may be small, the percentage error may be large when the amplitude of a mode is small. The significance of these repeatability errors is qualified by understanding how aberrations affect vision. Large repeatability errors, in a particular Zernike mode for example, could be considered insignificant if the effect of that aberration on visual performance is negligible. Further to the system-induced errors we measured operator-induced errors by testing two operators with and without realigning the instrument between repeated measurements. We found that repeatability errors were larger when the instrument was realigned but there was no significant difference between the operators. Next we studied subject-induced errors by changing the

pupil size and focus error of the model eye. This showed that third order aberrations were particularly affected by changing pupil radius and that all higher-order aberrations were affected by changing focus error. Low order aberrations were reasonably robust to changes in either. We also showed that repeatability errors can be minimised by recalculating Zernike coefficients over a fixed pupil radius. Measurements were then taken using human subjects to test for diurnal variations. We found that measured Zernike coefficients did not change significantly from day to day, although we cannot infer from this that our subjects' aberrations did not change since variations may have been within the limits of the instrument's repeatability. There was, however, a significant effect of the time of day that was evoked by the change in aberrations in only one eye. We suggest that the variations over even a small time period may be significant in some subjects, although we again cannot rule out the possibility of these variations being caused by factors not relating to real changes in the wavefront. It is important to take many repeated measurements immediately before performing psychophysical experiments to ensure that we have a representation of their ocular aberrations during the task that is as accurate as possible.

Contrast threshold for letter identification

Next we discussed two psychophysical experiments that tested the effects of defocus, coma and secondary astigmatism on letter-based visual tasks. In the first of these we measured how the contrast threshold for letter identification is affected by increasing amplitudes of these aberrations. We found that the increase in contrast threshold was greatest for secondary astigmatism, followed by defocus. Coma caused only a small increase in contrast threshold over the amplitudes we have tested. We suggested that this is because coma leaves the form of letters relatively intact, as predicted by our cross-correlation analysis. Interestingly, when we compared the increase in contrast threshold against our confusion metric we found that all three aberrations exhibited a similar relationship. This suggests that single letter identification can be predicted based on cross-correlation. Coma did not increase contrast threshold as much as defocus or secondary astigmatism but it also had a

lower confusion value. If the amplitude of coma had been increased such that its confusion value was similar to the highest amplitudes of defocus and secondary astigmatism, perhaps the decrease in performance would have also been similar.

Reading performance

The second psychophysical experiment tested the effect of these aberrations on reading performance. As with the letter identification task, the greatest impairment was caused by secondary astigmatism followed by defocus and coma was the least detrimental. We quantified this impairment in terms of the increase in average fixation duration over a sentence and in terms of the increase in the number of fixations made over that sentence. In addition to these global measures we analysed local, word-specific effects by looking at the fixations on a designated high or low lexical frequency target word. Recall that high lexical frequency words are easier to identify than low lexical frequency words and therefore we expect subjects to exhibit a shorter fixation on them (Rayner *et al.*, 2005; Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Rayner, 1998). For the total fixation time on the target word, the size of the lexical frequency effect was modulated by the amplitude of defocus and secondary astigmatism but not by the amplitude of coma. This means that lexical identification was increasingly affected by defocus and secondary astigmatism as their amplitude increased. This implies an affect on top-down processes between the word, letter and feature identification levels. Features normally present in letters are disrupted by phase changes that hinder their efficient identification and this disturbance propagates through to word identification. In the case of coma, top-down processes appear unaffected since lexical identification is unhindered by increasing amounts of aberration. This suggests that once the initial 'clean-up' of the retinal image has occurred and features and letters are extracted, word identification continues in the same manner as in the control condition. This difference between the effects of coma and those of defocus and secondary astigmatism are reflected in the confusion analysis. We showed that coma did not exhibit the same relationship between our confusion metric and the increase in average fixation duration as that exhibited

by defocus and secondary astigmatism. The confusability values are lower for coma over the amplitude range we have used but the increase in fixation duration is relatively higher than we would predict for defocus or secondary astigmatism. We have suggested that this 'extra' performance loss was due to lateral masking effects between words and letters within words. This occurs either because the signal from one letter overlaps with, and reduces the contrast of, the adjacent letter resulting in ordinary masking effects, or because it interferes with the signal integration process, resulting in crowding effects (Pelli, Palomares, & Majaj, 2004). These effects are analogous to a drop in the signal-to-noise ratio of that letter, which impedes feature and letter detection. Additionally, the 'smearing' effect of coma may cause interference in saccade planning by directing gaze to a less than optimal position in the word.

8.2. Future work

In light of the results presented in this thesis we suggest three topics for related future work.

1. Testing the accuracy of the Zywave aberrometer

In Chapter 5 we presented the results of several tests examining the repeatability of the Zywave aberrometer. Although repeatability is an important requirement of an instrument that measures ocular wavefront we also wish to know how accurate those measurements are. Our hope was to test this using customised phase plates that could be mounted inside the model eye. These phase plates could be used to compare the accuracy of each Zernike mode measured with the Zywave with the same measurements made with an interferometer, for example, in the absence of a scattering surface. Unfortunately this was not possible during the course of this work but we would suggest this is the next logical step in characterising this instrument. A method for creating such phase plates has previously been suggested by Rodríguez *et al.* (2006) and been shown to highlight small difference between measurements made with an interferometer and single pass measurements with a Shack-Hartmann sensor.

2. Testing the lateral masking hypothesis

We have shown in this thesis that when reading text there is a difference in the source of performance loss between coma and defocus or secondary astigmatism. We believe this performance loss is due to lateral masking effects and possibly disruptions to saccade planning. The hypothesis could be further tested by repeating our experiment with different orientations of coma. For example, inducing vertical rather than horizontal coma will cause the same amount of image degradation without introducing such strong masking effects. We expect that comparing letter identification and reading performance for these two orientations of coma would reveal an effect on the relationship between our confusion metric and the disruption to reading performance for horizontal coma only. It may also be possible to disentangle effects due to lateral masking from those due to saccade programming by using an eye tracker with a higher accuracy and by taking more repeated measurements of the same target word. Although the visual span is symmetric about the point of fixation the perceptual span is biased to the right (e.g. McConkie & Rayner, 1976). It may also be interesting to test the symmetry of these masking and saccade planning effects by inducing both positive and negative coma (i.e. comparing a 'smear' to the left with a 'smear' to the right).

3. Improved modelling of letter recognition

We have used a cross-correlation model to predict the confusability of single letters. While this model relates well to the experimental results it may not be intuitively the most appropriate. Cross-correlation simply compares two images with a transform of position and does not take account of stretching and skewing effects (see Figure 8.1 for example). A better way to include these differences in our confusion analysis might be to use another method of comparing letters such as using features rather than whole letter templates. A different approach might be to use mutual information (Shannon, 1948), which is measure of the amount of information one random variable contains about another. This technique

is typically used for image registration, for example when registering images taken by two different methods such as computed tomography (CT) and magnetic resonance imaging (MRI). We can define a transform such that one image maps on to the other in such a way as to maximise their joint entropy. Each image has a spatial intensity distribution and we can define a joint probability distribution which shows how those intensity distributions are correlated. This joint probability distribution is a 2-D histogram of corresponding pixel values between the two images. The histogram is built up by counting the number of times that a pair of pixels (one from each image) have the corresponding intensities. Example histograms are given in Figure 8.2. The image is considered to be registered when the dispersion (entropy) of the joint intensity distribution is at a minimum. In an analogy to this example, the distortions in the image of a letter caused by an aberration might be considered to be akin to the structure exhibited by the MR image. We want to compare this to an unaberrated letter that does not such distortions, just as the CT image does not have the structure exhibited by the MR image. An optimum transformation of one image, such as stretching, so that it has the maximum mutual information with the other image can be found.

Figure. 8.1: Miss-match in scale of unaberrated ‘n’ and aberrated ‘m’ that would not be accounted for by cross-correlation.

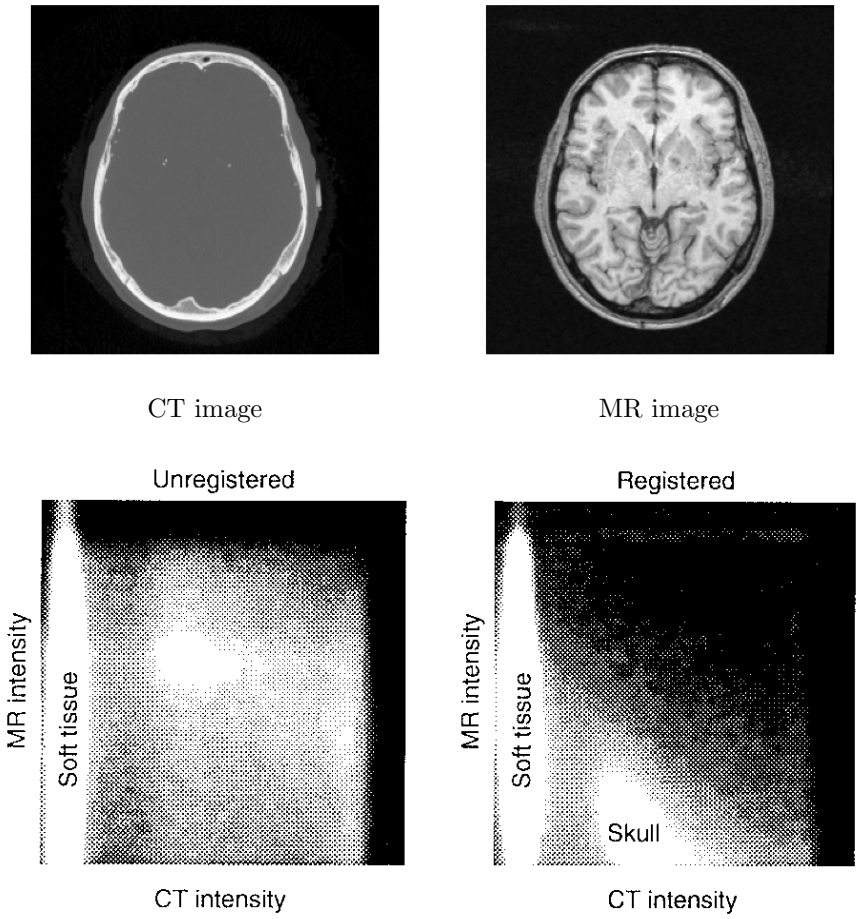

Figure. 8.2: Joint intensity histograms of a CT image and MRI image of the head showing the case when a) the images are not registered and b) the images are correctly registered (adapted from Maes *et al.*, 1997). Each point in the 2-D histogram is a count of the number of overlapping pixel pairs that have the corresponding intensities (CT intensity in one pixel and MR intensity in the other). If the images were identical the joint intensity distribution would be concentrated along the diagonal. In this example areas in the MR image have a range of intensity values for the centre of the head compared to the a uniform intensity in the CT image. The images are registered when the dispersion (entropy) of the joint intensity distribution is at a minimum.

Appendix 1: Analysis of variance

In this section we explain the statistical method of analysis of variance (ANOVA), which is used throughout this thesis. An ANOVA is a statistical test used to determine if the means of several groups of observations are significantly different, taking their variance into account. Where the samples are grouped by a factor, for example samples of some visual performance measure made with three different types of aberration, we can infer significant effects of the factor (in this case the type of aberration) on the performance measure. The two-way ANOVA tests observations grouped by two factors for three null hypotheses: that the means of observations grouped by one factor are the same, the means of observations grouped by the other factor are the same and that there is no interaction between factors (i.e. the effect of one factor does not vary with the effect of the other factor. For example, if we test three types of aberration at three different amplitudes of aberration we have two factors (type and amplitude of aberration) and an interaction would suggest that the effect of the amplitude on performance is different across the different types of aberration. Similarly the effect of the type of aberration on visual performance may be different at different amplitudes of aberration. If a significant interaction is shown then one-way ANOVAs can be performed on each factor separately to look for simple effects (e.g. to look for an effect of the amplitude of aberration separately for each type of aberration). If simple effects are

found then these can be broken down further by post-hoc pairwise comparisons. In the case where there are three groups (as is the case in this work) these comparisons equate to paired t-tests via Fisher's least significant difference (Howell, 1992, pp.356). These comparisons indicate at what level within a group the significant difference was found (e.g. between the first and second amplitude of given type of aberration).

The ANOVA tests for statistical significant by comparing the F-test statisitc (F) to the the F-distribution. The mean square of the observations for each factor, the mean square of the interaction and the mean square for the variation within each combination of factors are calculated. The test statistic, F is given by

$$F = \frac{MSE}{MSE_{Within}}, \tag{8.1}$$

where MSE is one of the mean squares listed above and MSE_{Within} is the mean square calculated within groups. If this F-test statistic is higher that the the value of the F-distribution for the given number of degrees of freedom in the data and the required significance level (usually 5%), then the effect is considered significant. There are three assumptions made by the ANOVA:

- **Independence:** The groups must be independent of one another.
- **Normality:** The data must be normally distributed (i.e. the frequency histogram of the data should follow a normal distrubtion).
- **Sphericity:** The variance of the difference scores between each pair of groups must the same.

If sphericity is violated corrections a must be applied. In our analysis we have used a Greenhouse-Geisser correction or a Huynh-Feldt correction depending on the correction factors calculated by Mauchly's test (which checks the sphericity of the data samples by comparing their covariance matrix). These corrections adjust the number of degrees of freedom to account for differences in the variances of the groups. Corrected degrees of freedom

are used when comparing the F-test statistic and these are quoted, where applicable in our analyses.

References

Aggarwala, K. R., Nowbotsing, S., & Kruger, P. B. (1995). Accommodation to monochromatic and white-light targets. *Investigative Ophthalmology and Visual Science*, *23*, 61-62.

Akutsu, H., Bedell, H. E., & Patel, S. S. (2000). Recognition threhsolds for letters with simulated dioptric blur. *Optometry and Vision Science*, *77*, 524-530.

Applegate, R. A., Ballentine, C., Gross, H., Sarver, E. J., & Sarver, C. (2003). Visual acuity as a function of Zernike mode and level of root mean square error. *Optometry and Vision Science*, *80*, 97-105.

Applegate, R. A., Marsack, J. D., Ramos, R., & Sarver, E. J. (2003). Interaction between aberrations to improve of reduce visual performance. *Journal of Cataract and Refractive Surgery*, *29*, 1487-1495.

Applegate, R. A., Sarver, E. J., & Khemsara, V. (2002). Are All Aberrations Equal? *Journal of Refractive Surgery*, *18*, S556-S562.

Arathorn, D. W., Yang, Q., Vogel, C. R., Zhang, Y., Tiruveedhula, P., & Roorda, A. (2007). Retinally stabilised cone-targeted stimulus delivery. *Optics Express*, *15*, 13731-13744.

Arnulf, A., & Dupuy, O. (1960). La transmission des contrastes par le système optique de l'œil et les seuilles des contrastes rétiens. *Comptes Rendues de l'Académie des Sciences*, *250*, 2757-2759.

Artal, P., Berrio, E., Guirao, A., & Piers, P. (2002). Contribution of the cornea and internal surfaces to the change in ocular aberrations with age. *Journal of the Optical Society*

of America A, *19*, 137-143.

Artal, P., Chen, L., Fernández, E. J., Singer, B., Manzanera, S., & Williams, D. R. (2004). Neural compensation for the eye's optical aberrations. *Journal of Vision*, *4*, 281-287.

Artal, P., Fernández, E. J., & Manzanera, S. (2002). Are optical aberrations during accommodation a significant problem for refractive surgery? *Journal of Refractive Surgery*, *18*, 563-566.

Artal, P., Guirao, A., Berrio, E., & Williams, D. R. (2001). Compensation of corneal aberrations by the internal optics in the human eye. *Journal of Vision*, *1*, 1-8.

Artal, P., Manzanera, S., Piers, P., & Weeber, H. (2010). Visual effect of the combined correction of spherical and longitudinal chromatic aberrations. *Optics Express*, *18*, 1637-1648.

Atchison, D. A., Guo, H., Charman, W. N., & Fisher, S. W. (2009). Blur limits for defocus, astigmatism and trefoil. *Vision Research*, *49*, 2393-2403.

Atchison, D. A., Guo, H., & Fisher, S. W. (2009). Limits of spherical blur determined with an adaptive optics system. *Ophthalmic and Physiological Optics*, *29*, 300-311.

Atchison, D. A., & Scott, D. H. (2002). Monochromatic aberrations of the human eye in horizontal visual field. *Journal of the Optical Society of America A*, *19*, 2180-2184.

Babcock, H. W. (1953). The possibility of compensating astronomical seeing. *Publications of the Astronomical Society of rhe Pacific*, *65*, 229-236.

Barnden, R. (1974). Calculation of axial polychromatic optical transfer function. *Optica Acta*, *21*, 981-1003.

Bennett, P. J., & Banks, M. S. (1987). Sensitivity loss among odd-symmetric mechanisms underlies phase anomalies in peripheral vision. *Nature*, *326*, 873-876.

Berny, F., & Slansky, S. (1969). Wavefront determination resulting from foucault test as applied to the human eye and visual instruments. In J. H. Dickson (Ed.), *Optical instruments and techniques* (p. 375-386). Oriel Press, Newcastle, Pa.

Berson, D. M., Dunn, F. A., & Takao, M. (2002). Phototransduction by retinal ganglion cells that set the circadian clock. *Science*, *295*, 1070-1073.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, *203*, 237-260.

Braddick, O. (1981). Is spatial phase degraded in peripheral vision and visual pathology? *Documenta Ophthalmologica Proceedings Series*, *30*, 255-262.

Bradshaw, G. (2005). Multimedia textbooks and student learning. *MERLOT Journal of Online Learning and Teaching*, *1*.

Bullimore, M. A., Dobos, M. J., & Twa, M. D. (2003). Repeatability and validity of refractive error using the bausch and lomb zywave. *Investigative Ophthalmology and Visual Science*, *44*, E-Abstract 4081.

Burakgazi, A. Z., Tinio, B., Bababyan, A., Niksarli, K. K., & Absell, P. (2006). Higher order aberrations in normal eyes measured with three different aberrometers. *Journal of Refractive Surgery*, *22*, 898-903.

Burr, D. C., Morrone, M. C., & Spinelli, D. (1989). Evidence for edge and bar detectors in human vision. *Vision Research*, *29*, 419-431.

Campbell, F. W., & Green, D. G. (1965). Optical and retinal factors affecting visual resolution. *Journal of Physiology*, *181*, 576-593.

Campbell, F. W., & Robson, J. G. (1964). Application of fourier analysis to the modulation response of the eye. *Journal of the Optical Society of America Society of America*, *54*, 581A.

Campbell, F. W., & Robson, J. G. (1968). Application of fourier analysis to the visibility of gratings. *Journal of Physiology (London)*, *197*, 551-566.

Campbell, F. W., & Westheimer, G. (1959). Factors influencing accommodation responses of the human eye. *Journal of the Optical Society of America A*, *49*, 568-571.

Cervino, A., Hosking, S. L., & Dunne, M. C. M. (2007). Operator-induced error in hartmann-shack wavefront sensing: Model eye study. *Journal of Cataract and Refractive Surgery*, *33*, 115-121.

Charman, W. N., & Chateau, N. (2003). The prospects for super-acuity: Limits to visual

performance after correction of monochromatic ocular aberration. *Ophthalmic and Physiological Optics*, *23*, 479-493.

Cheng, H., Barnett, J. K., Vilupuru, A. S., Marsack, J. D., Kasthurirangan, S., Applegate, R. A., *et al.* (2004). A population study on changes in wave aberrations with accommodation. *Journal of Vision*, *4*, 272-280.

Cheng, X., Himebaugh, N. L., Kollbaum, P. S., Thibos, L. N., & Bradley, A. (2003). Validation of a clinical shack-hartmann aberrometer. *Optometry and Vision Science*, *80*, 587-959.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Attention and performance vi. In S. Dornic (Ed.), (p. 535-555). Hillsdale, NJ: Erlbaum.

Dacey, D. M., Peterson, B. B., Robinson, F. R., & Gamlin, P. D. (2003). Fireworks in the primate retina: in vitro photodynamics reveals diverse lgn-projecting ganglion cell types. *Neuron*, *37*, 15-27.

Dalimier, E., Dainty, C., & Barbur, J. L. (2008). Effects of higher-order aberrations on contrast acuity as a function of light level. *Journal of Modern Optics*, *55*, 792-803.

DenBuurman, R., Boersma, T., & Gerrissen, J. F. (1981). Eye movements and the perceptual span in reading. *Reading Research Quarterly*, *16*, 227-235.

Diaz-Santana, L., & Dainty, J. C. (2001). Effects of retinal scattering in the ocular double-pass process. *Journal of the Optical Society of America A*, *18*, 1437-1444.

Diaz-Santana, L., & Dainty, J. C. (2003). Benefit of higher closed-loop bandwidths in ocular adaptive optics. *Optics Express*, *11*, 2597-2605.

Dobos, M. J., Twa, M. D., & Bullimore, M. A. (2009). An evaluation of the bausch and lomb zywave aberrometer. *Clinical and Experimental Optometry*, *92*, 238-245.

Dubra, A., Paterson, C., & Dainty, C. (2004). Study of the tear topography dynamics using a lateral shearing interferometer. *Optics Express*, *12*, 6278-6288.

Duchowski, A. (2003). *Eye tracking methodology: Theory and practice.* Springer-Verlag, New York.

Enroth-Cugell, C., & Robson, J. G. (1984). Functional characterisitics and diversity of cat

retinal ganglion cells. *Investigative Ophthalmology and Visual Science*, *25*, 250-267.

Fang, L., Wang, Y., & He, X. (2009). Evaluation of optical quality in white light from wavefront aberrations for a myopic population of human eyes. *Clinical and Experimental Optometry*, *92*, 313-319.

Fernández, E. J., Iglesias, I., & Artal, P. (2001). Closed-loop adaptive optics in the human eye. *Optics Letters*, *26*, 746-748.

Fernández, E. J., Prieto, P. M., & Artal, P. (2009). Binocular adaptive optics simulator. *Optics Letters*, *34*, 2628-2630.

Field, D. J., & Nachmias, J. (1984). Phase reversal discrimination. *Vision Research*, *24*, 333-340.

Findlay, J. M. (1978). Estimates on probability function: A mmore viurlent pest. *Perception & Psychophysics*, *23*, 181-185.

Fisher, D. F. (1976). Spatial factors in reading and search: The case for space. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes* (p. 417-427). L. Erlbaum.

Francis, N. W., & Kučera, H. (1985). Frequency analysis of english usage: Lexicon and grammar. *Journal of English Linguistics*, *18*, 64-70.

Frisby, J. P. (1980). *Seeing: Illusion, brain and mind*. Oxford University Press.

Gobbe, M., & Guillon, M. (2005). Corneal wavefront aberration measurements to detect keratoconus patients. *Contact Lens & Anterior Eye*, *28*, 57-66.

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: a comparison of single-channel and multi-channel models. *Vision Research*, *11*, 251-259.

Guirao, A., Porter, J., Williams, D. R., & Cox, I. G. (2002). Calculated impact of higher-order monochromatic aberrations on retinal image quality in a population of human eyes. *Journal of the Optical Society of America A*, *19*, 1-9.

Guirao, A., Redondo, M., & Artal, P. (2000). Optical aberrations of the human cornea as a function of age. *Journal of the Optical Society of America A*, *17*, 1697-1702.

Hall, J. L. (1968). Maximum-likelihood sequential procedure for estimation of psychometric functions. *Journal of the Acoustical Society of America*, *44*, 370.

Hament, W. J., Nabar, V. A., & Nuijts, R. M. M. A. (2002). Repeatability and validity of zywave aberrometer measurements. *Journal of Cataract and Refractive Surgery*, *28*, 2135-2141.

Hampson, K. M., Munro, I., Paterson, C., & Dainty, C. (2005). Weak correlation between the aberrations dynamics of the human eye and the cardiopulmonary system. *Journal of the Optical Society of America A*, *22*, 1241-1250.

Han, W., Kwan, W., Wang, J., Yip, S. P., & Yap, M. (2007). Influence of eyelid position on wavefront aberrations. *Ophthalmic and Physiological Optics*, *27*, 66-75.

Harvey, L. O. (1986). Efficient estimation of sensory thresholds. *Behaviour Research Methods, Instruments and Computers*, *18*, 623-632.

Harvey, L. O. (1997). Efficient estimation of sensory thresholds with ml-pest. *Spatial Vision*, *11*, 121-128.

Heath, G. G. (1956). The influence of visual acuity on accommodative responses of the eye. *American Journal of Optometry*, *33*, 513-524.

Hofer, H., Artal, P., Singer, B., Aragón, J. L., & Williams, D. R. (2001). Dynamics of the eye's wave aberration. *Journal of the Optical Society of America A*, *18*, 497-506.

Howell, D. C. (1992). *Statistical methods for psychology* (3rd edition ed.). CA: Duxbury Press.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics*, *40*, 431-439.

Jiménez, J. R., Ortiz, C., Hita, E., & Soler, M. (2008). Correlation between image quality and visual performance. *Journal of Modern Optics*, *55*, 783-790.

Jones, M. N., & Mewhort, D. J. K. (2004). Case-sensitive letter and bigram frequency counts from large-scale english corpora. *Behaviour Research Methods, Instruments and Computers*, *36*, 388-396.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions.

Nature, *290*, 91-97.

Kapoula, Z., Robinson, D., & Hain, T. (1986). Motion of the eye immediately after a saccade. *Experimental Brain Research*, *61*, 386-394.

Kepler, J. (1604). *Ad vitellionem paralipomena, quibus astronomiae pars optica traditur* (C. Marne & H. J. Auber, Eds.).

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & Psychophysics*, *63*, 1421-1455.

Koenderink, J. J., Bouman, M. A., Bueno de Mesquita, A. E., & Slappendel, S. (1978). Perimetry of contrast detection thresholds of moving spatial sine wave patterns. iv. the influence of the mean retinal illuminance. *Journal of the Optical Society of America*, *68*, 860-865.

Koh, S., Maeda, N., & Kuroda, T. (2002). Effect of tear film break-up on higher-order aberrations measured with a wavefront sensor. *American Journal of Ophthalmology*, *134*, 115-117.

Kruger, P. B., Mathews, S., Aggarwala, K. R., & Sanchez, N. (1993). Chromatic aberration and ocular focus: Fincham revisited. *Vision Research*, *33*, 1397-1411.

Kruger, P. B., & Pola, J. (1986). Stimuli for accommodation: Blur, chromatic aberration and size. *Vision Research*, *26*, 957-971.

Legge, G. E. (2007). *Psychophysics of reading in normal and low vision.* Lawrence Erlbaum Associates Inc.

Li, J., Xiong, Y., Wang, N., Li, S., Dai, Y., Xue, L., *et al.* (2009). Effects of spherical aberration on visual acuity at different contrasts. *Journal of Cataract and Refractive Surgery*, *35*, 1389-1395.

Li, S., Xiong, Y., Li, J., Wang, N., Dai, Y., Xue, L., *et al.* (2009). Effects of monochromatic aberration on visual acuity using adaptive optics. *Optometry and Vision Science*, *86*, 868-874.

Liang, C. L., Hank, S. H., & Chang, C. J. (2005). Comparison of higher-order wavefront aberrations with 3 aberrometers. *Journal of Cataract and Refractive Surgery*, *31*,

2153-2156.

Liang, J., Grimm, B., Goelz, S., & Bille, J. F. (1994). Objective measurement of wave aberrations of the human eye with the use of a hartmann-shack wave-front sensor. *Journal of the Optical Society of America A*, *11*, 1949-1957.

Liang, J., & Williams, D. R. (1997). Aberrations and retinal image quality of the normal human eye. *Journal of the Optical Society of America A*, *14*, 2873-2883.

Liang, J., Williams, D. R., & Miller, D. T. (1997). Supernormal vision and high-resolution retinal imaging through adaptive optics. *Journal of the Optical Society of America A*, *14*, 2884-2892.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science*, *4*, 6-14.

Liversedge, S. P., White, S. J., Findlay, J. M., & Rayner, K. (2006). Binocular coordination of eye movements during reading. *Vision Research*, *46*, 2363-2374.

López-Gil, N., Chateau, N., Castejón-Monchón, J. F., Artal, P., & Benito, A. (2003). Correcting ocular aberrations by soft contact lenses. *The South African Optomoetrist*, *62*, 173-177.

López-Gil, N., Fernández-Sánchez, V., Legras, R., Montés-Micó, R., Lara, F., & Nguyen-Khoa, J. L. (2008). Accommodation-related changes in monochromatic aberrations of the human eye as a function of age. *Investigative Ophthalmology and Visual Science*, *49*, 1736-1743.

Lu, F., Wu, J., Shen, Y., Qu, J., Wang, Q., Xu, C., *et al.* (2008). On the compensation of horizontal coma aberrations in young human eyes. *Ophthalmic and Physiological Optics*, *28*, 277-282.

Lundström, L., Manzanera, S., Prieto, P. M., & Ayala, D. B. (2007). Effect of optical correction and remaining aberrations on peripheral resolution acuity in the human eye. *Optics Express*, *15*, 12654-12661.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., & Suetens, P. (1997). Multimodality image registration by maximization of mutual information. *IEEE Transactions on*

Medical Imaging, *16*, 187-198.

Majaj, N. J., Pelli, D. G., Kurshan, P., & Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Research*, *42*, 1165-1184.

Malt, B. C., & Seamon, J. G. (1978). Peripheral and cognitive components of eye guidance in filled-space reading. *Perception & Psychophysics*, *23*, 399-402.

Mannos, J. L., & Sakrison, D. J. (1974). The effects of a visual fidelity criterion on the encoding of images. *IEEE Transactions on Information Theory*(525-535).

Mansfield, J. S., Legge, G. E., & Bane, M. C. (1996). Psychophysics of reading xv: Font effects in normal and low vision. *Investigative Ophthalmology and Visual Science*, *37*, 1492-2015.

Marcos, S., Burns, S. A., Moreno-Barriusop, E., & Navarro, R. (1999). A new approach to the study of ocular chromatic aberrations. *Vision Research*, *39*, 4309-4323.

Marcos, S., Sawides, L., Gambra, E., & Dorronsoro, C. (2008). Influence of adaptive optics ocular aberration correction on visual acuity at different luminances and contrast polarities. *Journal of Vision*, *8*, 1-12.

Marsack, J. D., Thibos, L. N., & Applegate, R. A. (2004). Metrics of optical quality derived from wave aberrations predict visual performance. *Journal of Vision*, *4*, 322-328.

Marčelja, S. (1980). Mathematical description of the responses of simple cortical cells. *Journal of the Acoustical Society of America*, *70*, 1297-1300.

Matthews, M. L. (1987). The influence of colour on ctr reading performance and subjective discomfort under operational conditions. *Applied Ergonomics*, *18*, 323-328.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. an account of basic findings. *Psychological Review*, *88*, 375-407.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: 1. the location of initial eye fixations on words. *Vision Research*, *28*, 1107-1118.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation

in reading. *Perception & Psychophysics*, *17*, 578-586.

McConkie, G. W., & Rayner, K. (1976). Assymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society*, *8*, 365-368.

McLellan, J., Marcos, S., & Burns, S. (2001). Age-related changes in monochromatic wave aberrations in the human eye. *Investigative Ophthalmology and Visual ScienceOphthmology and Visual Science*, *42*, 1390-1395.

Mirshahi, A., Bürhen, J., Herhardt, D., & Kohnen, T. (2003). In vivo and in vitro repeatability os hartmann-shack aberrometry. *Journal of Cataract and Refractive Surgery*, *29*, 2295-2301.

Morris, R. K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 268-281.

Myers, G. A., & Stark, L. (1990). Topology of the near response triad. *Ophthalmic and Physiological Optics*, *10*, 175-181.

Nachmias, J., & Weber, A. (1975). Discrimination of simple and complex gratings. *Vision Research*, *15*, 217-223.

Navarro, R., Moreno, E., & Dorronsoro, C. (1998). Monochromatic aberrations and point-spread functions of the human eye across the visual field. *Journal of the Optical Society of America A*, *15*, 2522-2529.

Netto, M. V., Dupps, W., & Wilson, S. E. (2006). Wavefront-guided ablation: Evidence for efficacy compared to traditional ablation. *American Journal of Ophthalmology*, *141*, 360-368.

Ninomiya, S., Fujikado, T., Kuroda, T., Maeda, N., Tano, Y., Oshika, T., *et al.* (2002). Changes of ocular aberration with accommodation. *American Journal of Ophthalmology*, *134*, 924-926.

Nyquist, H. (1928). Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, *47*, 617-644.

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and

glissade detection in eyetracking data. *Behaviour Research Methods*, *42*, 188-204.

Ogboso, Y. U., & Bedell, H. E. (1987). Magnitude of lateral chrmatic aberration across the retina of the human eye. *Journal of the Optical Society of America A*, *4*, 1666-1672.

Oshika, T., Klyce, S. D., Applegate, R. A., Howland, H. C., & El Danasoury, M. A. (1999). Comparison of corneal wavefront aberrations after photorefractive keratectomy and laser in situ keratomileusis. *American Journal of Ophthalmology*, *127*, 1-7.

Oshika, T., Okamoto, C., Samejima, T., Tokunaga, T., & Miyata, K. (2006). Contrast sensitivity function and ocular higher-order aberrations in normal human eyes. *Ophthalmology*, *113*, 1807-1812.

Osterberg, G. (1935). Topography of the layer of rods and cones in the human retina. *Acta Ophthalmologica*, *6*, 1-103.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*, 1136-1169.

Pelli, D. G., Robson, J. G., & wilkins, A. J. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences*, *2*, 187-199.

Pentland, A. (1980). Maximum likelihood estimation: The best pest. *Perception & Psychophysics*, *28*, 377-379.

Pollatsek, A., & Rayner, K. (1982). Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 817-833.

Prins, N., & Kingdom, F. A. A. (2009). *Palamedes: Matlab routines for analyzing psychophysical data.* Available from `http://www.palamedestoolbox.org`

Ravikumar, S., Bradley, A., & Thibos, L. (2010). Phase changes induced by optical aberrations degrade letter and face acuity. *Journal of Vision*, *18*, 1-12.

Ravikumar, S., Thibos, L. N., & Bradley, A. (2008). Calculation of retinal image quality for polychromatic light. *Journal of the Optical Society of America A*, *25*, 2395-2407.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*, 65-81.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422.

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science*, *206*, 468-469.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: effects of word frequency, verb complexity and lexical ambiguity. *Memory & Cognition*, *14*, 191-201.

Rayner, K., Inhoff, A. W., Morrison, R., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 167-179.

Rayner, K., Liversedge, S. P., & White, S. J. (2005). Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research*, *46*, 310-323.

Rayner, K., & McConkie, G. (1976). What guides a reader's eye movements? *Vision Research*, *16*, 829-837.

Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics*, *27*, 537-544.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of the stimulus material. *Journal of Experimental Psychology*, *81*, 274-280.

Rentschler, I., & Treutwien, B. (1985). Loss of spatial phase relationships in extrafoveal vision. *Nature*, *313*, 308-310.

Rocha, K. M., Vabre, L., Chateau, N., & Krueger, R. R. (2010). Enhanced visual acuity and image perception following correction of highly aberrated eyes using an adaptive optics visual simulator. *Journal of Refractive Surgery*, *26*, 52-56.

Rocha, K. M., Vabre, L., Harms, F., Chateau, N., & Krueger, R. R. (2007). Effects of zernike wavefront aberrations on visual acuity measured using electromagnetic adptive optics technology. *Journal of Refractive Surgery*, *23*, 953-959.

Rodríguez, P., Navarro, R., Arines, J., & Bará, S. (2006). A new calibration set of phase plates for ocular aberrometers. *Journal of Refractive Surgery*, *22*, 275-284.

Roorda, A. (2011). Adaptive optics for studying visual function: A comprehensive review. *Journal of Vision*, *11*, 1-21.

Rouger, H., Benard, Y., & Legras, R. (2009). Effects of monochromatic induced aberrations on visual performance measured by adaptive optics technology. *Journal of Refractive Surgery*, *26*, 578-587.

Rovamo, J., Virsu, V., & Näsänen, R. (1978). Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision. *Nature*, *271*, 54-56.

Savitzky, A., & Golay, M. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*, 1627-1639.

Sawides, L., Gambra, E., Pascual, D., Dorronsoro, C., & Marcos, S. (2010). Visual performance with real-life tasks under adaptive-optics ocular aberration correction. *Journal of Vision*, *10*, 1-12.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379-423.

Smirnov, M. S. (1961). Measurement of the wave aberration of the human eye. *Biofizika*, *6*, 687-703.

Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature*, *369*, 395-397.

Spragins, A. B., Lefton, L. A., & Fisher, D. F. (1976). Eye movements while reading and searching spatially transformed text: A developmental examination. *Memory & Cognition*, *4*, 36-42.

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Science*, *5*, 156-163.

Taylor, M. M., & Creelman, C. D. (1967). Pest: Efficient stimates on probability functions. *Journal of the Acoustical Society of America*, *41*, 782-787.

Thibos, L. N., Applegate, R. A., Schwiegerling, J. T., & Webb, R. (2000). Standards for reporting the optical aberrations of eyes. In *Vision science and its applications, osa technical digest (optical society of america, 2000)* (p. paper SuC1).

Thibos, L. N., Hong, X., Bradley, A., & Applegate, R. A. (2004). Accuracy and precision of objective refraction from wavefront aberrations. *Journal of Vision*, *4*, 329-351.

Thibos, L. N., Hong, X., Bradley, A., & Cheng, X. (2002). Statistical variation of aberration structure and image quality in a normal population of healthy eyes. *Journal of the Optical Society of America A*, *19*, 2329-2348.

Thorn, F., & Schwartz, F. (1990). Effects of dioptric blur on snellen and grating acuity. *Optometry and Vision Science*, *67*, 3-7.

Townsend, J. T., Taylor, S. G., & Brown, D. R. (1971). Lateral masking for letters with unlimited viewing time. *Perception & Psychophysics*, 375-378.

Underwood, N. R., & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly*, *20*, 153-162.

Van Essen, D. C. (2004). Organization of visual areas in macaque and human cerebral cortex. In L. Chalupa & J. S. Werner (Eds.), (p. 507-521). MIT Press.

Villa, C., Gutiérrez, R., Jiménez, J. R., & González-Méijome, J. M. (2007). Night vision disturbances after successful lasik surgery. *British Journal of Ophthalmology*, *91*, 1031-1037.

von Helmholtz, H. (1867). *Helmholtz's treatise on physiological optics.* Translated by J. P. C. Southall (1924).

Wallington, E., & Symons, P. (2008). Video eyetracker toolbox user manual [Computer software manual]. Available from `http://www.crsltd.com/`

Watson, A. B., & Pelli, D. G. (1983). Quest: A bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*, 113-120.

Weber, R., & Daroff, R. (1972). Corrective movements following refixation saccades: Type and control systems. *Vision Research*, *12*, 467-475.

Westheimer, G. (1960). Modulation thresholds for sinusoidal light distributions on the retina. *Journal of Physiology*, *152*, 67-74.

White, S. J., & Liversedge, S. P. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology*, *16*, 52-78.

Williams, D. R., Artal, P., Navarro, R., McMahon, M. J., & Brainard, D. H. (1996). Off-axis optical quality and retinal sampling in the human eye. *Vision Research*, *36*, 1103-1114.

Williams, D. R., Yoon, G. Y., Porter, J., Guirao, A., Hofer, H., & Cox, I. (2000). Visual benefit of correcting higer order aberrations of the eye. *Journal of Refractive Surgery*, *16*, S554-S559.

Yoon, G. Y., Jeong, T. M., Cox, I. G., & Williams, D. R. (2004). Vision improvement by correcting higer-order aberrations with phase plates in normal eyes. *Journal of Refractive Surgery*, *20*, S523-S527.

Yoon, G.-Y., & Williams, D. R. (2002). Visual performance after correcting the monochromatic and chromatic aberrations of the eye. *Journal of the Optical Society of America A*, *19*, 266-275.

Young, T. (1801). On the mechanism of the eye. *Philosophical Transactions of the Royal Society of London*, *91*, 23-28.

Zernike, F. (1934). Diffraction theory of knife-edge test and its improved form, phase contrast method. *Monthly Notices of the Royal Astronomical Society*, *94*, 377-384.